

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 10:35:46 ; Search time 51.0566 seconds
(without alignments)
358,688 Million cell updates/sec

Title: US-10-144-679-1

Perfect score: 33

Sequence: 1 cactctcttcgagccggctcgaaatagtgagt 33

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodaca/2/ina/5A-COMB.seq:*
 - 2: /cgn2_6/ptodaca/2/ina/5B-COMB.seq:*
 - 3: /cgn2_6/ptodaca/2/ina/6A-COMB.seq:*
 - 4: /cgn2_6/ptodaca/2/ina/6B-COMB.seq:*
 - 5: /cgn2_6/ptodaca/2/ina/PTCUS-COMB.seq:*
 - 6: /cgn2_6/ptodaca/2/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	60.6	1122	4	US-09-489-039A-2435
C 2	20	60.6	536185	4	US-09-214-808-1
C 3	19	57.6	1221	4	US-09-252-991A-4504
C 4	18.8	57.0	50	3	US-09-010-641-35
C 5	18.8	57.0	50	3	US-09-356-281-35
C 6	18.8	57.0	3182	4	US-09-221-017B-1029
C 7	18.6	56.4	1722	4	US-09-434-840-1
C 8	18.6	56.4	1722	4	US-09-434-840-3
C 9	18.6	56.4	1731	4	US-09-434-840-62
C 10	18.6	56.4	1732	4	US-09-434-840-54
C 11	18.6	56.4	1732	4	US-09-434-840-56
C 12	18.6	56.4	1732	4	US-09-434-840-60
C 13	18.6	56.4	1733	4	US-09-434-840-58
C 14	18.6	56.4	7213	4	US-09-634-238-20
C 15	18.6	56.4	11168	4	US-09-434-840-5
C 16	18.4	55.8	1236	4	US-09-252-991A-1087
C 17	18.4	55.8	1947	4	US-09-252-991A-1048
C 18	18.4	55.8	2636	1	US-08-554-612C-12
C 19	18.4	55.8	2898	1	US-08-554-612C-51
C 20	18.4	55.8	2809	1	US-08-554-612C-10
C 21	18.4	55.8	2809	1	US-08-554-612C-11
C 22	18.4	55.8	5843	1	US-08-554-612C-2
C 23	18.2	55.2	699	4	US-09-489-039A-1572
C 24	18	54.5	528	3	US-09-489-039A-3807
C 25	18	54.5	1035	3	US-08-191-160-4
C 26	18	54.5	1176	2	US-08-537-811-41
C 27	18	54.5	1770	3	US-08-191-160-3

C 28 18 54.5 2220 2 US-08-932-176A-3
C 29 18 54.5 3750 3 US-08-181-160-22
C 30 18 54.5 5410 4 US-09-221-017B-70
C 31 18 54.5 3472 4 US-08-150-204E-96
C 32 18 54.5 10399 4 US-08-961-527-160
C 33 18 54.5 46899 1 US-08-471-119A-1
C 34 17.8 53.9 47 4 US-09-641-638-1285
C 35 17.8 53.9 201 4 US-09-621-976-296
C 36 17.8 53.9 1000 4 US-09-641-638-631
C 37 17.8 53.9 1251 2 US-09-211-930-2
C 38 17.8 53.9 1251 3 US-09-340-993-2
C 39 17.8 53.9 1251 3 US-09-468-442-2
C 40 17.8 53.9 1153 2 US-09-211-930-8
C 41 17.8 53.9 1353 3 US-09-340-993-8
C 42 17.8 53.9 1353 4 US-09-468-442-8
C 43 17.8 53.9 1482 4 US-09-489-039A-6520
C 44 17.8 53.9 1542 4 US-09-345-473E-13
C 45 17.8 53.9 2343 4 US-09-641-638-652

ALIGNMENTS

RESULT 1
US-09-489-039A-2435/C
Sequence 2435, Application US/09489039A
Patent No. 6610836

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 2435
LENGTH: 1122
TYPE: DNA

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-2435

Query Match 60.6%, Score 20, DB 4, Length 1122;
Best Local Similarity 82.1%, Pred. No. 9,
Matches 23, Conservative 0, Mismatches 5, Indels 0, Gaps 0;

Cy 5 TCTTCTCCGACCGGTCGAAATAGTGAG 32
Db 1103 TCTTCTCCGACCGGTCGAAATAGTGAG 1076

RESULT 2

US-09-214-808-1

Sequence 1, Application US/09214808A

Patent No. 6475791

GENERAL INFORMATION:

APPLICANT: Rosenthal, Andre

APPLICANT: Freiberg, Christoph

APPLICANT: Perret, Xavier Philippe

APPLICANT: Broughton, William John

TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic

Patent No. 6475793

TITLE OF INVENTION: plasmid

FILE REFERENCE: CARP0068

CURRENT APPLICATION NUMBER: US/09/214,808A

CURRENT FILING DATE: 1999-06-22

PRIOR APPLICATION NUMBER: ECT/IB97/00950

PRIOR FILING DATE: 1997-07-10

NUMBER OF SEQ ID NOS: 1

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 536165

```

/ TYPE: DNA
/ ORGANISM: Rhizobium
US-09-214-808-1

Query Match      60.6%; Score 20; DB 4; Length 536165;
Best Local Similarity 82.1%; Pred. No. 21;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 CTCCTCTCCGAGCCGCTCGAATAGTCA 31
Db 455880 CTCCTCACCAGCCGCTCGACCTGCTCA 455907

RESULT 3
US-09-252-991A-4504/c
; Sequence 4504, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,780
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4504
; LENGTH: 1221
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4504

Query Match      57.6%; Score 19; DB 4; Length 1221;
Best Local Similarity 81.5%; Pred. No. 25;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CATCTCTCTCCGAGCCGCTCGAATA 27
Db 1128 CAGCACTTCGCGAGCATGTCCGAATA 1102

RESULT 4
US-09-010-641-35/c
; Sequence 35, Application US/09010641
; Patent No. 6121023
; GENERAL INFORMATION:
; APPLICANT: ROMANO, JOSEPH W.
; APPLICANT: SHUTLIFF, ROXANNE
; APPLICANT: WILLIAMS, KIMBERLY G.
; TITLE OF INVENTION: ISOTHERMAL AMPLIFICATION BASED ASSAY FOR
; TITLE OF INVENTION: THE DETECTION AND QUANTIFICATION OF CHEMOKINES RANTES,
; TITLE OF INVENTION: MIP-1ALPHA AND MIP-1BETA
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AZZO NOBEL PATENT DEPARTMENT
; STREET: 1300 PICCARD DRIVE, SUITE 206
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/010,641
; FILING DATE: 22-JAN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

```

/ NAME: KLESNER, SHARON N.
/ REGISTRATION NUMBER: 36,335
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 301-948-7400
/ TELEFAX: 301-948-9751
/ INFORMATION FOR SEQ ID NO: 35:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 50 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
US-09-010-641-35

Query Match      57.0%; Score 18.8; DB 3; Length 50;
Best Local Similarity 76.7%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Oy 4 CTCCTCCGAGCCGCTCGAATAGTCA 33
Db 42 CTCCTCCGAGCCGCTCGAATAGTCA 13

RESULT 5
US-09-356-281-35/c
; Sequence 35, Application US/09356281
; Patent No. 6218154
; GENERAL INFORMATION:
; APPLICANT: ROMANO, JOSEPH W.
; APPLICANT: SHUTLIFF, ROXANNE
; APPLICANT: WILLIAMS, KIMBERLY G.
; TITLE OF INVENTION: ISOTHERMAL AMPLIFICATION BASED ASSAY FOR
; TITLE OF INVENTION: THE DETECTION AND QUANTIFICATION OF CHEMOKINES RANTES,
; TITLE OF INVENTION: MIP-1ALPHA AND MIP-1BETA
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AZZO NOBEL PATENT DEPARTMENT
; STREET: 1300 PICCARD DRIVE, SUITE 206
; CITY: ROCKVILLE
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/356,281
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/010,641
; FILING DATE: 22-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, SHARON N.
; REGISTRATION NUMBER: 36,335
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-948-7400
; TELEFAX: 301-948-9751
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 50 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-356-281-35

Query Match      57.0%; Score 18.8; DB 3; Length 50;
Best Local Similarity 76.7%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

QY 4 CTCCTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 42 CTCCTCTCCGAGCCGGTCGAAATAGTGAGT 13

RESULT 6

US-09-221-0178-1029/c
; Sequence 1029, Application US/092210178
; Patent No. 6444729
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,0178
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Mohrroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.C0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1029:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3182 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1...3182
US-09-221-0178-1029

Query Match 57.0%; Score 18.8; DB 4; Length 3182;
Best Local Similarity 76.7%; Pred. No. 35;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGTG 30
DB 1334 CAGCTCTCTCTCCGAGCCGGTCGAAATAGTG 1305

RESULT 7

US-09-434-840-1
; Sequence 1, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jiraseg, Dayadevi
; APPLICANT: Toole, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)..(1616)
US-09-434-840-1

Query Match 56.4%; Score 18.6; DB 4; Length 1722;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 1155 CAACCTCTTAGCCGAGCCCACTCGACATTCGGAAT 1187

RESULT 8

US-09-434-840-3
; Sequence 3, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jiraseg, Dayadevi
; APPLICANT: Toole, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (36)..(1064)
US-09-434-840-3

Query Match 56.4%; Score 18.6; DB 4; Length 1722;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 1155 CAACCTCTTAGCCGAGCCCACTCGACATTCGGAAT 1187

```
RESULT 9
US-09-434-840-62
; Sequence 56, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootle, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1548)
US-09-434-840-62

Query Match          56.4%; Score 18.6; DB 4; Length 1731;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 1165 CAATCTTTAGCCGAGCCACTCGACATTGCGAAT 1197

RESULT 10
US-09-434-840-54
; Sequence 54, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootle, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1626)
; OTHER INFORMATION: nucleotide sequence of wild type PAD4 cDNA cloned
; OTHER INFORMATION: in pCR1.1
US-09-434-840-54

Query Match          56.4%; Score 18.6; DB 4; Length 1732;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 1165 CAATCTTTAGCCGAGCCACTCGACATTGCGAAT 1197
```

```
RESULT 11
US-09-434-840-56
; Sequence 56, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootle, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 56
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1074)
US-09-434-840-56

Query Match          56.4%; Score 18.6; DB 4; Length 1732;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 1165 CAATCTTTAGCCGAGCCACTCGACATTGCGAAT 1197

RESULT 12
US-09-434-840-60
; Sequence 60, Application US/09434840
; Patent No. 6620985
; GENERAL INFORMATION:
; APPLICANT: Glazebrook, Jane
; APPLICANT: Jirage, Dayadevi
; APPLICANT: Tootle, Tina L
; APPLICANT: Zhou, Nan
; APPLICANT: Feys, Bart
; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
; FILE REFERENCE: 043503.0009
; CURRENT APPLICATION NUMBER: US/09/434,840
; CURRENT FILING DATE: 1999-11-04
; EARLIER APPLICATION NUMBER: 09/190,733
; EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 1732
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1158)
US-09-434-840-60

Query Match          56.4%; Score 18.6; DB 4; Length 1732;
Best Local Similarity 72.7%; Pred. No. 40;
Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCCGGTCGAAATAGTGAGT 33
DB 1165 CAATCTTTAGCCGAGCCACTCGACATTGCGAAT 1197
```


RESULT 13
 US-09-434-840-58
 ; Sequence 58, Application US/09434840
 ; Patent No. 6620985
 ; GENERAL INFORMATION:
 ; APPLICANT: Glaxobrook, Jane
 ; APPLICANT: Jiraga, Dayadevi
 ; APPLICANT: Tootle, Tina L
 ; APPLICANT: Zhou, Nan
 ; APPLICANT: Feys, Bart
 ; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
 ; FILE REFERENCE: 043503.0009
 ; CURRENT APPLICATION NUMBER: US/09/434,840
 ; CURRENT FILING DATE: 1999-11-04
 ; EARLIER APPLICATION NUMBER: 09/190,733
 ; EARLIER FILING DATE: 1998-11-12
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 58
 ; LENGTH: 1733
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(546)
 US-09-434-840-58

Query Match 56.4%; Score 18.6; DB 4; Length 1733;
 Best Local Similarity 72.7%; Pred. No. 40;
 Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 CATCTCTTCCGAGCGCTCGAATAGTCACT 33
 DB 1166 CAATCTTACCGAGCGCTCGACATTCGCAAT 1198

RESULT 14
 US-09-634-238-20
 ; Sequence 20, Application US/09634238
 ; Patent No. 6544772
 ; GENERAL INFORMATION:
 ; APPLICANT: Glenn, Matthew
 ; APPLICANT: Havukkale, Ilkka J.
 ; APPLICANT: Bloksberg, Leonard, N.
 ; APPLICANT: Lubbers, Mark W.
 ; APPLICANT: Dekker, James
 ; APPLICANT: Christenson, Anna C.
 ; APPLICANT: Holland, Ross
 ; APPLICANT: O'Toole, Paul W.
 ; APPLICANT: Reid, Julian R.
 ; APPLICANT: Coolbeax, Timothy
 ; TITLE OF INVENTION: Polynucleotides, materials incorporating
 ; TITLE OF INVENTION: them and methods for using them.
 ; FILE REFERENCE: 11006.104301
 ; CURRENT APPLICATION NUMBER: US/09/634,238
 ; CURRENT FILING DATE: 2000-08-08
 ; NUMBER OF SEQ ID NOS: 422
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 20
 ; LENGTH: 7213
 ; TYPE: DNA
 ; ORGANISM: Lactobacillus rhamnosus
 US-09-634-238-20

Query Match 56.4%; Score 18.6; DB 4; Length 7213;
 Best Local Similarity 72.7%; Pred. No. 49;
 Matches 24; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 1 CATCTCTTCCGAGCGCTCGAATAGTCACT 33
 DB 2733 CATCTCTTCCGAGCGCTCGAATAGTCAAT 2765

RESULT 15
 US-09-434-840-5
 ; Sequence 5, Application US/09434840
 ; Patent No. 6620985
 ; GENERAL INFORMATION:
 ; APPLICANT: Glaxobrook, Jane
 ; APPLICANT: Jiraga, Dayadevi
 ; APPLICANT: Tootle, Tina L
 ; APPLICANT: Zhou, Nan
 ; APPLICANT: Feys, Bart
 ; TITLE OF INVENTION: PAD4 COMPOSITIONS AND METHODS THEREFOR
 ; FILE REFERENCE: 043503.0009
 ; CURRENT APPLICATION NUMBER: US/09/434,840
 ; CURRENT FILING DATE: 1999-11-04
 ; EARLIER APPLICATION NUMBER: 09/190,733
 ; EARLIER FILING DATE: 1998-11-12
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1168
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (7327)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (7423)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8753)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8755)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8768)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8774)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8776)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8784)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8796)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (8799)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (9030)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (9325)
 ; OTHER INFORMATION: n = g or a or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (9466)

REGISTRATION NUMBER: 36,677
 REFERENCE/DOCKET NUMBER: 06519/004001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 322-5070
 TELEFAX: (415) 854-0875
 INFORMATION FOR SEQ ID NO: 12:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2636 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-554-612C-12

Query Match 55.8%; Score 18.4; DB 1; Length 2636;
 Best Local Similarity 78.6%; Pred. No. 52;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATCTCTCTCCGAGCCGTCGAATAGT 29
 Db 1606 ATGCTCTCTCCAGCCCGTCAACATATT 1633

RESULT 19
 US-08-554-612C-51
 Sequence 51, Application US/08554612C
 Patent No. 5747660
 GENERAL INFORMATION:
 APPLICANT: Orlicky, David
 TITLE OF INVENTION: PROSTAGLANDIN F2' RECEPTOR REGULATORY
 TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
 NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESSES:
 ADDRESSES: Fish & Richardson P.C.
 STREET: 2200 Sand Hill Road, Suite 100
 CITY: Menlo Park
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 94025

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/554,612C
 FILING DATE: No. 5747660ember 6, 1995
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
 NAME: Sherwood, Pamela
 REGISTRATION NUMBER: 36,677
 REFERENCE/DOCKET NUMBER: 06519/004001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 322-5070
 TELEFAX: (415) 854-0875
 INFORMATION FOR SEQ ID NO: 51:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2898 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 US-08-554-612C-51

Query Match 55.8%; Score 18.4; DB 1; Length 2898;
 Best Local Similarity 78.6%; Pred. No. 53;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATCTCTCTCCGAGCCGTCGAATAGT 29
 Db 1595 ATGCTCTCTCCAGCCCGTCAACATATT 1622

RESULT 20
 US-08-554-612C-10
 Sequence 10, Application US/08554612C
 Patent No. 5747660
 GENERAL INFORMATION:
 APPLICANT: Orlicky, David
 TITLE OF INVENTION: PROSTAGLANDIN F2' RECEPTOR REGULATORY
 TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
 NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESSES:
 ADDRESSES: Fish & Richardson P.C.
 STREET: 2200 Sand Hill Road, Suite 100
 CITY: Menlo Park
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 94025

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/554,612C
 FILING DATE: No. 5747660ember 6, 1995
 CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:
 NAME: Sherwood, Pamela
 REGISTRATION NUMBER: 36,677
 REFERENCE/DOCKET NUMBER: 06519/004001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 322-5070
 TELEFAX: (415) 854-0875
 INFORMATION FOR SEQ ID NO: 10:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2909 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 US-08-554-612C-10

Query Match 55.8%; Score 18.4; DB 1; Length 2909;
 Best Local Similarity 78.6%; Pred. No. 53;
 Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATCTCTCTCCGAGCCGTCGAATAGT 29
 Db 1606 ATGCTCTCTCCAGCCCGTCAACATATT 1633

RESULT 21
 US-08-554-612C-11
 Sequence 11, Application US/08554612C
 Patent No. 5747660
 GENERAL INFORMATION:
 APPLICANT: Orlicky, David
 TITLE OF INVENTION: PROSTAGLANDIN F2' RECEPTOR REGULATORY
 TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
 NUMBER OF SEQUENCES: 51
 CORRESPONDENCE ADDRESSES:
 ADDRESSES: Fish & Richardson P.C.
 STREET: 2200 Sand Hill Road, Suite 100
 CITY: Menlo Park
 STATE: California
 COUNTRY: U.S.A.
 ZIP: 94025

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/554,612C

```
/ FILING DATE: No. 5747660ember 6, 1995
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sherwood, Pamela
/ REGISTRATION NUMBER: 36,677
/ REFERENCE/DOCKET NUMBER: 06519/004001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 322-5070
/ TELEFAX: (415) 854-0875
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2909 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
US-08-554-612C-11

Query Match 55.8%; Score 18.4; DB 1; Length 2909;
Best Local Similarity 78.6%; Pred. No. 53;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATCTCTTCTCCGAGCCGTCGAAATAGT 29
DB 1606 ATGTTCTTCCAGCCGTCACATATT 1633

RESULT 22
US-08-554-612C-2
/ Sequence 2, Application US/08554612C
/ Patent No. 5747660
/ GENERAL INFORMATION:
/ APPLICANT: Orlicky, David
/ TITLE OF INVENTION: PROSTAGLANDIN P2 RECEPTOR REGULATORY
/ TITLE OF INVENTION: PROTEIN AND THERAPEUTIC USES
/ NUMBER OF SEQUENCES: 51
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson P.C.
/ STREET: 2200 Sand Hill Road, Suite 100
/ CITY: Menlo Park
/ STATE: California
/ COUNTRY: U.S.A.
/ ZIP: 94025
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ FILING DATE: No. 5747660ember 6, 1995
/ APPLICATION NUMBER: US/08/554,612C
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Sherwood, Pamela
/ REGISTRATION NUMBER: 36,677
/ REFERENCE/DOCKET NUMBER: 06519/004001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 322-5070
/ TELEFAX: (415) 854-0875
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 5843 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
US-08-554-612C-2

Query Match 55.8%; Score 18.4; DB 1; Length 5843;
Best Local Similarity 78.6%; Pred. No. 58;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATCTCTTCTCCGAGCCGTCGAAATAGT 29
```

```
DB 1644 ATGTTCTTCCAGCCGTCACATATT 1671

RESULT 23
US-09-489-039A-1572/c
/ Sequence 1572, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 1572
/ LENGTH: 699
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1572

Query Match 55.2%; Score 18.2; DB 4; Length 699;
Best Local Similarity 74.2%; Pred. No. 53;
Matches 23; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGTCGAAATAGTCA 31
DB 309 CAGCTGTTCCAGCTCCCGTCTGCTATAGCCA 279

RESULT 24
US-09-489-039A-3807
/ Sequence 3807, Application US/09489039A
/ Patent No. 6610836
/ GENERAL INFORMATION:
/ APPLICANT: Gary Breton et. al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
/ TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 2709.2004001
/ CURRENT APPLICATION NUMBER: US/09/489,039A
/ CURRENT FILING DATE: 2000-01-27
/ PRIOR APPLICATION NUMBER: US 60/117,747
/ PRIOR FILING DATE: 1999-01-29
/ NUMBER OF SEQ ID NOS: 14342
/ SEQ ID NO 3807
/ LENGTH: 528
/ TYPE: DNA
/ ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3807

Query Match 54.5%; Score 18; DB 4; Length 528;
Best Local Similarity 80.8%; Pred. No. 62;
Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 8 TCTCCGAGCCGTCGAAATAGTCAAGT 33
DB 68 TCCGCGAGCCGTCGAAAGAGTGAAT 93

RESULT 25
US-08-191-160-4/c
/ Sequence 4, Application US/08191160
/ Patent No. 6210675
/ GENERAL INFORMATION:
/ APPLICANT: Highfield, Peter Edmund
/ APPLICANT: Rodgers, Brian Colin
/ APPLICANT: Tedder, Richard Seron
/ APPLICANT: Barbara, John Anthony James
/ TITLE OF INVENTION: Viral Agent
/ NUMBER OF SEQUENCES: 25
```

CORRESPONDENCE ADDRESS:
ADDRESSES: Rothwell, Pigg, Ernst & Kurz
STREET: 1700 K Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 KB storage
COMPUTER: IBM AT compatible
OPERATING SYSTEM: MS-DOS V3.2
SOFTWARE: Wordperfect 5.0 (DOS text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191,160
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 07/628,516
FILING DATE: 17 DEC 1990
APPLICATION NUMBER: UK 89 28 562.1
FILING DATE: 18 DEC 1989
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 414.0
FILING DATE: 27 FEB 1990
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 814.1
FILING DATE: 03 MAR 1990
ATTORNEY/AGENT INFORMATION:
NAME: E. Anthony Flegg
REGISTRATION NUMBER: 27,195
REFERENCE/DOCKET NUMBER: 1645-103A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 833-5740
TELEFAX: (202) 833-5744
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1035 base pairs
TYPE: nucleotide with corresponding protein
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to genomic RNA
ORIGINAL SOURCE:
ORGANISM: human; serum infectious for PT-NANBH
IMMEDIATE SOURCE: clone JG3 from cDNA library in lambda gtl1
FEATURE:
LOCATION: from 1 to 1035 bp portion of the PT-NANBH
LOCATION: polyprotein
OTHER INFORMATION: probably encodes viral non-structural
OTHER INFORMATION: proteins
US-08-191-160-4

```

Query Match          54.5%; Score 18; DB 3; Length 1035;
Best Local Similarity 80.8%; Pred. No. 68;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4  CTCTCTCCGAGCGCGTCGAAATAGT 29
          |||||
DB       420  CTCCGCTCGAGCGCGTCGAAAGAGT 395

```

RESULT 26
US-08-537-811-41/c
Sequence 41, Application US/0853781.1
Patent No. 5910405
GENERAL INFORMATION:
APPLICANT: CHO, JOONG YOUNG
APPLICANT: CHOI, BEONG YOUNG
APPLICANT: KIM, CHUN HYUNG
APPLICANT: SO, HONG SEOB
APPLICANT: YANG, JAE YOUNG
APPLICANT: KIM, IN SOO
APPLICANT: KIM, JOO HO

TITLE OF INVENTION; IMPROVED HCV DIAGNOSTIC
TITLE OF INVENTION; AGENTS
NUMBER OF SEQUENCES; 48
CORRESPONDENCE ADDRESS;
ADDRESSEE; Pennie & Edmonds LLP
STREET; 1155 Avenue of the Americas
CITY; New York
STATE; NY
COUNTRY; USA
ZIP; 10036-2711
COMPUTER READABLE FORM;
MEDIUM TYPE; Diskette
COMPUTER; IBM Compatible
OPERATING SYSTEM; DOS
SOFTWARE; FASTSEQ Version 2.0
CURRENT APPLICATION DATA;
APPLICATION NUMBER; US/08/537,811
FILING DATE; 24-OCT-1995
CLASSIFICATION; 435
PRIOR APPLICATION DATA;
APPLICATION NUMBER; PCT/KR94/00040
FILING DATE; 29-APR-1994
APPLICATION NUMBER; KR 93-7440
FILING DATE; 30-APR-1993
ATTORNEY/AGENT INFORMATION;
NAME; Jones, Ili, Harry C
REGISTRATION NUMBER; 20,280
REFERENCE/DOCKET NUMBER; 8512-037-999
TELECOMMUNICATION INFORMATION;
TELEPHONE; 212-790-9090
TELEFAX; 212-869-9741
TELEX; 65141 PENNIE
INFORMATION FOR SEQ ID NO; 41:
SEQUENCE CHARACTERISTICS;
LENGTH; 1176 base pairs
TYPE; nucleic acid
STRANDEDNESS; single
TOPOLOGY; linear
MOLECULE TYPE; cDNA
FEATURE;
OTHER INFORMATION; KHCV NGS-1.2, Fig. 3
US-08-537-811-41

```

Query Match      54.5%; Score 18; DB 2; Length 1176;
Best Local Similarity 80.8%; Pred. NO. 70;
Matches 21; Conservative 0; Mismatches 5; Indels

QY  4  CTCCTTCGCGAGCGCGTCCGAATAGT  29
      |||||
DB   480  CTCGCGCTCGGAGCGCGTCCGAACAGT  455

```

```

RESULT 27.
US-08-191-160-3/c
; Sequence 3, Application US/08191160
; Patent No. 6210675
; GENERAL INFORMATION:
; APPLICANT: Highfield, Peter Edmund
; APPLICANT: Rodgers, Brian Colin
; APPLICANT: Tedder, Richard Seton
; APPLICANT: Barbara, John Anthony James
; TITLE OF INVENTION: Viral Agent
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurtz
; STREET: 1700 K Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 KB storage
; COMPUTER: IBM AT compatible

```

```

/ OPERATING SYSTEM: MS-DOS V3.2
/ SOFTWARE: Wordperfect 5.0 (DOS text)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/191.160
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/628.516
/ FILING DATE: 17 DEC 1990
/ APPLICATION NUMBER: UK 89 28 562.1
/ FILING DATE: 18 DEC 1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: UK 90 04 414.0
/ FILING DATE: 27 FEB 1990
/ APPLICATION NUMBER: UK 90 04 814.1
/ FILING DATE: 03 MAR 1990
/ ATTORNEY/AGENT INFORMATION:
/ NAME: E. Anthony Piggy
/ REGISTRATION NUMBER: 27,195
/ REFERENCE/DOCKET NUMBER: 1645-103A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 833-5740
/ TELEFAX: (202) 833-5744
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1770 base pairs
/ TYPE: nucleotide with corresponding protein
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA to genomic RNA
/ ORIGINAL SOURCE:
/ ORGANISM: human, serum infectious for PT-NANBH
/ IMMEDIATE SOURCE:
/ LIBRARY: clone J62 from cDNA library in lambda gt11
/ FEATURE:
/ LOCATION: from 1 to 1770 bp portion of the PT-NANBH
/ LOCATION: polyprotein
/ OTHER INFORMATION: probably encodes viral non-structural
/ OTHER INFORMATION: proteins
/
US-08-191-160-3

Query Match 54.5%; Score 18; DB 3; Length 1770;
Best Local Similarity 80.8%; Pred. No. 74;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CTCCTCCGAGCCGGTCGAAATAGT 29
Db 135 CTCCTCCGAGCCGGTCGAAAGT 110

RESULT 28
US-08-932-376A-3/c
Sequence 3, Application US/08932376A
Patent No. 5869309
GENERAL INFORMATION:
APPLICANT: Politingo, Michael
APPLICANT: Tonsi, Sean M.
APPLICANT: Usher, John J.
APPLICANT: Burnett K, William V.
APPLICANT: Romancik, Guna
TITLE OF INVENTION: CEPHALOSPORIN ESTERASE GENE FROM
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: Bristol-Myers Squibb Company
STREET: Rt. 206 & Provinceline Road
CITY: Princeton
STATE: New Jersey
COUNTRY: USA
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

Query Match 54.5%; Score 18; DB 2; Length 2220;
Best Local Similarity 80.8%; Pred. No. 76;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 TTCTCCGAGCCGGTCGAAATAGT 32
Db 1344 TTCTCCGAGCCGGTCGACATGTCTAG 1319

RESULT 29
US-08-191-160-22/c
Sequence 22, Application US/08191160
Patent No. 6210675
GENERAL INFORMATION:
APPLICANT: Highfield, Peter Edmund
APPLICANT: Rodgers, Brian Colin
APPLICANT: Tedder, Richard Seton
APPLICANT: Barbara, John Anthony James
TITLE OF INVENTION: Viral Agent
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESS: Rothwell, Figg, Brnst & Kurz
STREET: 1700 K Street
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20006
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 KB storage
COMPUTER: IBM AT compatible
OPERATING SYSTEM: MS-DOS V3.2
SOFTWARE: Wordperfect 5.0 (DOS text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/191.160
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/628.516
FILING DATE: 17 DEC 1990
APPLICATION NUMBER: UK 89 28 562.1
FILING DATE: 18 DEC 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 414.0
FILING DATE: 27 FEB 1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 90 04 814.1
FILING DATE: 03 MAR 1990
ATTORNEY/AGENT INFORMATION:
NAME: E. Anthony Piggy
REGISTRATION NUMBER: 27,195

```

REFERENCE/DOCKET NUMBER: 1645-103A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 833-5740
 TELEFAX: (202) 833-5744
 INFORMATION FOR SEQ ID NO: 22:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3750 base pairs
 TYPE: nucleotide with corresponding protein
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to genomic RNA
 ORIGINAL SOURCE:
 ORGANISM: human; serum infectious for PT-NANBH
 IMMEDIATE SOURCE:
 LIBRARY: cDNA clones from 3' end of the genome
 FEATURE:
 LOCATION: from 1 to 3750 bp portion of the PT-NANBH
 LOCATION: polyprotein
 OTHER INFORMATION: viral non-structural proteins
 US-08-191-160-22

Query Match 54.5%; Score 18; DB 3; Length 3750;
 Best Local Similarity 80.8%; Pred. No. 82;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

US-08-191-160-22

Qy 4 CTCCTCTCGAGCCGGTCGAATAGT 29
 |||||
 Db 2115 CTCCTCTCGAGCCGGTCGAATAGT 2090

RESULT 30
 US-09-221-017B-70
 Sequence 70, Application US/09221017B
 Patent No. 644799
 GENERAL INFORMATION:
 APPLICANT: ROSS, Bruce C.
 TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
 NUMBER OF SEQUENCES: 1120
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FORSTER
 STREET: 755 PAGE MILL ROAD
 CITY: PALO ALTO
 STATE: CA
 COUNTRY: USA
 ZIP: 94304-1018
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: Windows
 SOFTWARE: FASTSEQ for Windows Version 2.0b
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/221,017B
 FILING DATE: 23-DEC-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P01182
 FILING DATE: 31-DEC-1997
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P01546
 FILING DATE: 30-JAN-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: P02911
 FILING DATE: 09-APR-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/AU98/01023
 FILING DATE: 10-DEC-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: MORROW, Gladys H.
 REGISTRATION NUMBER: 32,430
 REFERENCE/DOCKET NUMBER: 27340-20021.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 650-813-5600
 TELEFAX: 650-494-0792

TELEX: 706141
 INFORMATION FOR SEQ ID NO: 70:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5410 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: UNKNOWN
 ORIGINAL SOURCE:
 ORGANISM: PORPHYROMONAS GINGIVALIS
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: 1...5410
 US-09-221-017B-70

Query Match 54.5%; Score 18; DB 4; Length 5410;
 Best Local Similarity 80.8%; Pred. No. 86;
 Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 5 TCTTCTCGAGCCGGTCGAATAGT 30
 |||||
 Db 4186 TCTTCTCGAGCCGGTCGAATAGT 4211

RESULT 31
 US-08-150-204B-96/c
 Sequence 96, Application US/08150204B
 Patent No. 6538126
 GENERAL INFORMATION:
 APPLICANT: CHD, Joong Myung
 LEE, Yong Beom
 PARK, Young Woo
 LIM, Kook Jin
 CHOI, Deog Young
 SO, Hong Seob
 KIM, Chun Hyung
 KIM, Sung Taek
 YANG, Jae Young
 TITLE OF INVENTION: HEPATITIS C DIAGNOSTICS AND VACCINES
 NUMBER OF SEQUENCES: 128
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: YANG, Jae Young
 STREET: 386-1, Doryong-dong, Yuseong-gu
 CITY: Daejeon
 STATE: Daejeon
 COUNTRY: Republic of Korea
 ZIP: 305-340
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.5inch, 1.44MB storage
 COMPUTER: IBM PC/pentium
 OPERATING SYSTEM: Windows
 SOFTWARE: Microsoft Word
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/150,204B
 FILING DATE: 20-APR-1994
 CLASSIFICATION: <unknown>
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: KR 91-9510
 FILING DATE: 10-JUN-1991
 APPLICATION NUMBER: KR 91-13601
 FILING DATE: 6-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: SHAHAN ISLAM, Esq.
 REGISTRATION NUMBER: 32,507
 REFERENCE/DOCKET NUMBER: 2695/FLK
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 940-8564
 TELEFAX: (212) 940-8776
 INFORMATION FOR SEQ ID NO: 96
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9472 base pairs

```
/
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA
/ FEATURES:
/ OTHER INFORMATION: KHCY-L8C1, Fig. 2
/ SEQUENCE DESCRIPTION: SEQ ID NO: 96
US-08-150-2048-96

Query Match          54.5%; Score 18; DB 4; Length 9472;
Best Local Similarity 80.8%; Pred. No. 93;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      4 CTCTCTCCGAGCGGTCGGAATAGT 29
DB      7128 CTCCTCGAGCGGTCGGAATAGT 7103

RESULT 32
US-08-961-527-160/c
/ Sequence 160, Application US/08961527
/ Patent No. 6420135
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 391
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/961,527
/ FILING DATE:
/ CLASSIFICATION: 424
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brookes, A. Anders
/ REGISTRATION NUMBER: 36,373
/ REFERENCE/DOCKET NUMBER: PB340P1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (301) 309-8504
/ TELEFAX: (301) 309-8512
/ INFORMATION FOR SEQ ID NO: 160:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 10399 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
US-08-961-527-160

Query Match          54.5%; Score 18; DB 4; Length 10399;
Best Local Similarity 80.8%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      6 CTCTCCGAGCGGTCGGAATAGTGA 31
DB      1401 CTCTCCGAGCGGTCGGAATAGTGA 1376

RESULT 33
US-08-471-119A-1/c
/ Sequence 1, Application US/08471119A
/ Patent No. 5827706
```

```
/
/ GENERAL INFORMATION:
/ APPLICANT: Leitner, Brnet
/ APPLICANT: Schneider, Elisabeth
/ APPLICANT: Schoergerdorfer, Kurt
/ APPLICANT: Weber, Gerhard
/ TITLE OF INVENTION: Cyclosporin Synthetase
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: No. 5827706artis Corporation
/ STREET: 59 Route 10
/ CITY: East Hanover
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07936
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/471,119A
/ FILING DATE: 06-JUN-1995
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kasenoff, Melvyn
/ REGISTRATION NUMBER: 26,389
/ REFERENCE/DOCKET NUMBER: 100-8029/CONT/CONT
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201 503 8474
/ TELEFAX: 201 503 8474
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 46899 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHEICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Tolypocladium niveum
/ STRAIN: ATCC 34921
US-08-471-119A-1

Query Match          54.5%; Score 18; DB 1; Length 46899;
Best Local Similarity 80.8%; Pred. No. 1.2e+02;
Matches 21; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY      1 CACTCTCTCCGAGCGGTCGGAAT 26
DB      11867 CAGATCTCTCCAGCCGTCGATAT 11842

RESULT 34
US-09-641-638-1285
/ Sequence 1285, Application US/09641638
/ Patent No. 6432648
/ GENERAL INFORMATION:
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Bougueleret, Lydie
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Cohen, Annick
/ TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
/ TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
/ FILE REFERENCE: GENSET.051CP1
/ CURRENT APPLICATION NUMBER: US/09/641,638
/ CURRENT FILING DATE: 2000-08-16
/ PRIOR APPLICATION NUMBER: US 09/502,330
/ PRIOR FILING DATE: 2000-02-11
/ PRIOR APPLICATION NUMBER: US 60/133,200
/ PRIOR FILING DATE: 1999-05-07
/ PRIOR APPLICATION NUMBER: US 09/275,267
/ PRIOR FILING DATE: 1999-03-23
```



```

; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 1285
; LENGTH: 47
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 24
; OTHER INFORMATION: 10-507-170 : polymorphic base A or G
US-09-641-638-1285

Query Match      53.9%; Score 17.8; DB 4; Length 47;
Best Local Similarity 71.0%; Pred. No. 54;
Matches 22; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CACTCTCTCCGAGCCGTCGAATAGTGA 31
Db 10 CACCTCTCTACCAAGGAGTGCATAGTGA 40

RESULT 35
US-09-621-976-296/c
; Sequence 296, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S. Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 296
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 40..201
; NAME/KEY: sig_peptide
; LOCATION: 40..162
; OTHER INFORMATION: Von Heijne matrix
; OTHER INFORMATION: score 4.0999990463257
; OTHER INFORMATION: seq GILLMIIRLIFS/KT
US-09-621-976-296

Query Match      53.9%; Score 17.8; DB 4; Length 201;
Best Local Similarity 75.9%; Pred. No. 67;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 4 CTCTCTCCGAGCCGTCGAATAGTGA 32
Db 62 CACTGTCTCGTGGCGTCCCATAGTGA 14

RESULT 36
US-09-641-638-631
; Sequence 631, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641.638

```

```

; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 631
; LENGTH: 1000
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 501
; OTHER INFORMATION: 10-507-170 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 481..500
; OTHER INFORMATION: 10-507-170.misl, potential
; NAME/KEY: misc_binding
; LOCATION: 502..521
; OTHER INFORMATION: 10-507-170.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 332..350
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 739..758
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 489..513
; OTHER INFORMATION: 10-507-170 potential probe
US-09-641-638-631

Query Match      53.9%; Score 17.8; DB 4; Length 1000;
Best Local Similarity 71.0%; Pred. No. 83;
Matches 22; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CACTCTCTCCGAGCCGTCGAATAGTGA 31
Db 487 CACCTCTCTACCAAGGAGTGCATAGTGA 517

RESULT 37
US-09-211-930-2
; Sequence 2, Application US/09211930
; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tytell, R. No. 5962265r1e
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; TITLE OF INVENTION: SERINE/THREONINE KINASE
; FILE REFERENCE: PFM.70296
; CURRENT APPLICATION NUMBER: US/09/211.930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-211-930-2

Query Match      53.9%; Score 17.8; DB 2; Length 1251;
Best Local Similarity 75.9%; Pred. No. 86;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 5 TCTTCTCCGAGCCGTCGAATAGTGA 33

```

Db 327 TCTTCTCGAGCTGGTCCATTTCAGT 355
||||| ||||| ||||| ||||| |||||

RESULT 38
US-09-340-993-2
; Sequence 2, Application US/09340993
; Patent No. 6034228
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6034228ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM 70296.N1
; CURRENT APPLICATION NUMBER: US/09/340,993
; CURRENT FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0 & US 09/211,930
; EARLIER FILING DATE: 1997-12-19 & 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-340-993-2

Query Match 53.9%; Score 17.8; DB 3; Length 1251;
Best Local Similarity 75.9%; Pred. No. 86;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 TCTTCTCGAGCTGGTCCATTTCAGT 33
||||| ||||| ||||| ||||| |||||
Db 327 TCTTCTCGAGCTGGTCCATTTCAGT 355

RESULT 39
US-09-468-442-2
; Sequence 2, Application US/09468442
; Patent No. 6300098
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 6300098ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION SERINE/THREONINE KINASE
; FILE REFERENCE: PHM 70296.N1
; CURRENT APPLICATION NUMBER: US/09/468,442
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: US 09/340,993
; EARLIER FILING DATE: 1999-06-25
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; EARLIER APPLICATION NUMBER: US 09/211,930
; EARLIER FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1251
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-468-442-2

Query Match 53.9%; Score 17.8; DB 4; Length 1251;
Best Local Similarity 75.9%; Pred. No. 86;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 TCTTCTCGAGCTGGTCCATTTCAGT 33
||||| ||||| ||||| ||||| |||||
Db 327 TCTTCTCGAGCTGGTCCATTTCAGT 355

RESULT 40
US-09-211-930-8
; Sequence 8, Application US/09211930

; Patent No. 5962265
; GENERAL INFORMATION:
; APPLICANT: Tyrell E. No. 5962265ris
; APPLICANT: William Craig Moore
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN SIGNAL TRANSDUCTION
; FILE REFERENCE: PHM 70296
; CURRENT APPLICATION NUMBER: US/09/211,930
; CURRENT FILING DATE: 1998-12-15
; EARLIER APPLICATION NUMBER: GB 9726851.0
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 1353
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-211-930-8

Query Match 53.9%; Score 17.8; DB 2; Length 1353;
Best Local Similarity 75.9%; Pred. No. 87;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 5 TCTTCTCGAGCTGGTCCATTTCAGT 33
||||| ||||| ||||| ||||| |||||
Db 333 TCTTCTCGAGCTGGTCCATTTCAGT 361

Search completed: May 24, 2004, 12:31:17
JOB time : 53.0566 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 11:41:02 ; Search time 199.245 Seconds
(without alignments)
752.721 Million cell updates/sec

Title: US-10-144-679-1

Perfect score: 33

Sequence: 1 CATCTCTTCTCCGACCGGTGGAATAGTAGT 33

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	33	16	US-10-144-679-1
2	33	100.0	33	16	US-10-144-679-82
3	33	100.0	33	17	US-10-144-094-1
4	33	100.0	33	17	US-10-144-094-82
5	31.4	95.2	33	16	US-10-144-679-87
6	27.8	84.2	50	17	US-10-144-679-20
7	27.8	84.2	50	17	US-10-144-094-20
8	26.4	80.0	50	16	US-10-144-679-18
9	26.4	80.0	50	16	US-10-144-679-19
10	26.4	80.0	50	17	US-10-144-094-18
11	26.4	80.0	50	17	US-10-144-094-19
12	24	72.7	50	16	US-10-144-679-21
13	24	72.7	50	17	US-10-144-094-21
14	21.8	66.1	33	16	US-10-144-679-7

15	21.8	66.1	33	17	US-10-144-094-7
16	20	60.6	698	15	US-10-074-475-102
17	20	60.6	765	16	US-10-369-493-28205
18	20	60.6	771	16	US-10-369-493-30964
19	20	60.6	1018	9	US-09-974-300-1317
20	20	60.6	1217	15	US-10-074-475-103
21	20	60.6	536165	10	US-09-939-964-1
22	19.8	60.0	453	13	US-10-424-599-50726
23	19.8	60.0	2581	13	US-10-425-114-30662
24	19.2	58.2	34214	9	US-09-782-378A-27
25	19.2	58.2	50	16	US-10-144-679-14
26	19.2	58.2	50	17	US-10-144-679-1
27	19.2	58.2	9025608	15	US-10-156-761-1
28	19	57.6	444	9	US-09-796-692-4311
29	19	57.6	444	15	US-10-040-862-4311
30	19	57.6	444	16	US-10-057-475B-4311
31	19	57.6	444	16	US-10-154-884B-4311
32	18.8	57.0	51	16	US-10-144-679-17
33	18.8	57.0	51	17	US-10-144-094-17
34	18.8	57.0	220	13	US-10-424-599-74527
35	18.8	57.0	495	9	US-09-974-300-5658
36	18.8	57.0	846	13	US-10-282-122A-14316
37	18.8	57.0	1132	13	US-10-425-114-25452
38	18.8	57.0	1229	13	US-10-425-114-22309
39	18.8	57.0	1371	13	US-10-282-122A-14254
40	18.8	57.0	3182	13	US-10-194-163-1029
41	18.8	57.0	9025608	15	US-10-156-761-1
42	18.6	56.4	49	16	US-10-144-679-15
43	18.6	56.4	49	15	US-10-144-679-16
44	18.6	56.4	49	17	US-10-144-094-15
45	18.6	56.4	49	17	US-10-144-094-16

ALIGNMENTS

RESULT 1

US-10-144-679-1
; Sequence 1, Application US/10144679
; Publication No. US20030215810A1
; GENERAL INFORMATION:
; APPLICANT: LU, YI
; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
; TITLE OF INVENTION: COLOR CHANGES
; FILE REFERENCE: 9800240-0019
; CURRENT APPLICATION NUMBER: US/10/144.679
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Trans-cleaving
; OTHER INFORMATION: deoxyribozyme 17E
US-10-144-679-1

Query Match 100.0% Score 33; DB 16; Length 33;

Best Local Similarity 100.0%; Pred No. 8.9e+06; Mismatches 0; Indels 0; Gaps 0;

Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATCTCTTCTCCGACCGGTGGAATAGTAGT 33

Db 1 CATCTCTTCTCCGACCGGTGGAATAGTAGT 33

RESULT 2

US-10-144-679-82

; Sequence 82, Application US/10144679

; Publication No. US20030215810A1

; GENERAL INFORMATION:

```

; APPLICANT: LIU, YI
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
; TITLE OF INVENTION: COLOR CHANGES
; FILE REFERENCE: 9800240-0019
; CURRENT APPLICATION NUMBER: US/10/144,679
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic chimeric
; OTHER INFORMATION: substrate
; FEATURE:
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: Synthetic chimeric
; OTHER INFORMATION: substrate
US-10-144-679-82

```

```

Query Match 100.0%; Score 33; DB 16; Length 33;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33
Db 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33

RESULT 3
US-10-144-094-1
; Sequence 1, Application US/10144094
; Publication No. US20040023216A1
; GENERAL INFORMATION:
; APPLICANT: LIU, YI
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
; FILE REFERENCE: 10322/44
; CURRENT APPLICATION NUMBER: US/10/144,094
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Trans-cleaving
; OTHER INFORMATION: deoxyribozyme 178
US-10-144-094-1

```

```

Query Match 100.0%; Score 33; DB 17; Length 33;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33
Db 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33

RESULT 4
US-10-144-094-82
; Sequence 82, Application US/10144094
; Publication No. US20040023216A1
; GENERAL INFORMATION:
; APPLICANT: LIU, YI
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
; FILE REFERENCE: 10322/44
; CURRENT APPLICATION NUMBER: US/10/144,094
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 84

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 82
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic chimeric
; OTHER INFORMATION: substrate
; OTHER INFORMATION: Description of Combined DNA/RNA Molecule: Synthetic chimeric
; OTHER INFORMATION: substrate
US-10-144-094-82

```

```

Query Match 100.0%; Score 33; DB 17; Length 33;
Best Local Similarity 100.0%; Pred. No. 8.9e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33
Db 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33

```

```

RESULT 5
US-10-144-679-87
; Sequence 87, Application US/10144679
; Publication No. US20030215810A1
; GENERAL INFORMATION:
; APPLICANT: LIU, YI
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
; TITLE OF INVENTION: COLOR CHANGES
; FILE REFERENCE: 9800240-0019
; CURRENT APPLICATION NUMBER: US/10/144,679
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 33
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Inactive
; OTHER INFORMATION: deoxyribozyme 178-C
US-10-144-679-87

```

```

Query Match 95.2%; Score 31.4; DB 16; Length 33;
Best Local Similarity 97.6%; Pred. No. 5.2e-05;
Matches 32; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33
Db 1 CATCTCTCTCCGAGCGGTCGAAATAGTGAGT 33

```

```

RESULT 6
US-10-144-679-20
; Sequence 20, Application US/10144679
; Publication No. US20030215810A1
; GENERAL INFORMATION:
; APPLICANT: LIU, YI
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
; TITLE OF INVENTION: COLOR CHANGES
; FILE REFERENCE: 9800240-0019
; CURRENT APPLICATION NUMBER: US/10/144,679
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Zn-DNA
US-10-144-679-20

Query Match 84.2%; Score 27.8; DB 16; Length 50;
Best Local Similarity 93.5%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCTCTCTCCGAGCCGGTCGAAATAGTCAG 32
DB 1 ATCTCTCTCCGAGCCGGTCGAAATAGTCAG 31

RESULT 7

US-10-144-094-20

Sequence 20, Application US/10144094

Publication No. US20040023216A1

GENERAL INFORMATION:

APPLICANT: LU, YI

APPLICANT: LIU, JUWEN

TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR

FILE REFERENCE: 10322/44

CURRENT APPLICATION NUMBER: US/10/144,094

CURRENT FILING DATE: 2002-05-10

NUMBER OF SEQ ID NOS: 84

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 20

LENGTH: 50

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA

US-10-144-094-20

Query Match 84.2%; Score 27.8; DB 17; Length 50;
Best Local Similarity 93.5%; Pred. No. 0.0028;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCTCTCTCCGAGCCGGTCGAAATAGTCAG 32
DB 1 ATCTCTCTCCGAGCCGGTCGAAATAGTCAG 31

RESULT 8

US-10-144-679-18

Sequence 18, Application US/10144679

Publication No. US20030215810A1

GENERAL INFORMATION:

APPLICANT: LU, YI

APPLICANT: LIU, JUWEN

TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON

FILE REFERENCE: 9800240-0019

CURRENT APPLICATION NUMBER: US/10/144,679

CURRENT FILING DATE: 2002-05-10

NUMBER OF SEQ ID NOS: 88

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 18

LENGTH: 50

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Zn-DNA

US-10-144-679-18

Query Match 80.0%; Score 26.4; DB 16; Length 50;
Best Local Similarity 96.4%; Pred. No. 0.013;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCTCTCCGAGCCGGTCGAAATAGTCGA 31
DB 22 CTATCTCCGAGCCGGTCGAAATAGTCGA 49

RESULT 9

US-10-144-679-19

Sequence 19, Application US/10144679

Publication No. US20030215810A1

GENERAL INFORMATION:

APPLICANT: LU, YI

APPLICANT: LIU, JUWEN

TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON

FILE REFERENCE: 9800240-0019

CURRENT APPLICATION NUMBER: US/10/144,679

CURRENT FILING DATE: 2002-05-10

NUMBER OF SEQ ID NOS: 88

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 19

LENGTH: 50

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic

OTHER INFORMATION: Zn-DNA

US-10-144-679-19

Query Match 80.0%; Score 26.4; DB 16; Length 50;
Best Local Similarity 96.4%; Pred. No. 0.013;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCTCTCCGAGCCGGTCGAAATAGTCGA 31
DB 22 CTATCTCCGAGCCGGTCGAAATAGTCGA 49

RESULT 10

US-10-144-094-18

Sequence 18, Application US/10144094

Publication No. US20040023216A1

GENERAL INFORMATION:

APPLICANT: LU, YI

APPLICANT: LIU, JUWEN

TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR

FILE REFERENCE: 10322/44

CURRENT APPLICATION NUMBER: US/10/144,094

CURRENT FILING DATE: 2002-05-10

NUMBER OF SEQ ID NOS: 84

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 18

LENGTH: 50

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA

US-10-144-094-18

Query Match 80.0%; Score 26.4; DB 17; Length 50;
Best Local Similarity 96.4%; Pred. No. 0.013;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCTCTCCGAGCCGGTCGAAATAGTCGA 31
DB 22 CTATCTCCGAGCCGGTCGAAATAGTCGA 49

RESULT 11

US-10-144-094-19

Sequence 19, Application US/10144094

Publication No. US20040023216A1

GENERAL INFORMATION:

APPLICANT: LU, YI

APPLICANT: LIU, JUWEN

TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR

FILE REFERENCE: 10322/44

; CURRENT APPLICATION NUMBER: US/10/144,094
 ; CURRENT FILING DATE: 2002-05-10
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 19
 ; LENGTH: 50
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA
 US-10-144-094-19

Query Match 80.0%; Score 26.4; DB 17; Length 50;
 Best Local Similarity 96.4%; Pred. No. 0.013;
 Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 CTCTCTCCGAGCCGTCGAAATAGTGA 31
 DB 22 CTATCTCCGAGCCGTCGAAATAGTGA 49

RESULT 12
 US-10-144-679-21
 ; Sequence 21, Application US/10144679
 ; Publication No. US20030215810A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, JUEWEN
 ; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
 ; TITLE OF INVENTION: COLOR CHANGES
 ; FILE REFERENCE: 9800240-0019
 ; CURRENT APPLICATION NUMBER: US/10/144,679
 ; CURRENT FILING DATE: 2002-05-10
 ; NUMBER OF SEQ ID NOS: 88
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 50
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Zn-DNA
 US-10-144-679-21

Query Match 72.7%; Score 24; DB 16; Length 50;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CTCTCTCCGAGCCGTCGAAATAGT 29
 DB 6 CTCTCTCCGAGCCGTCGAAATAGT 29

RESULT 13
 US-10-144-094-21
 ; Sequence 21, Application US/10144094
 ; Publication No. US20040023216A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, JUEWEN
 ; TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
 ; FILE REFERENCE: 10322/44
 ; CURRENT APPLICATION NUMBER: US/10/144,094
 ; CURRENT FILING DATE: 2002-05-10
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 50
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA
 US-10-144-094-21

Query Match 72.7%; Score 24; DB 17; Length 50;
 Best Local Similarity 100.0%; Pred. No. 0.18;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 CTCTCTCCGAGCCGTCGAAATAGT 29
 DB 6 CTCTCTCCGAGCCGTCGAAATAGT 29

RESULT 14
 US-10-144-679-7
 ; Sequence 7, Application US/10144679
 ; Publication No. US20030215810A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, YI
 ; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
 ; TITLE OF INVENTION: COLOR CHANGES
 ; FILE REFERENCE: 9800240-0019
 ; CURRENT APPLICATION NUMBER: US/10/144,679
 ; CURRENT FILING DATE: 2002-05-10
 ; NUMBER OF SEQ ID NOS: 88
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 33
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Variant of
 ; OTHER INFORMATION: deoxyribozyme named 17E1
 US-10-144-679-7

Query Match 66.1%; Score 21.8; DB 16; Length 33;
 Best Local Similarity 78.8%; Pred. No. 2;
 Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGTCGAAATAGTGAGT 33
 DB 1 CATCTCTCTCTCCGAGCCGTCGAAATAGTGAGT 33

RESULT 15
 US-10-144-094-7
 ; Sequence 7, Application US/10144094
 ; Publication No. US20040023216A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LIU, YI
 ; TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
 ; FILE REFERENCE: 10322/44
 ; CURRENT APPLICATION NUMBER: US/10/144,094
 ; CURRENT FILING DATE: 2002-05-10
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 33
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Variant of
 ; OTHER INFORMATION: deoxyribozyme named 17E1
 US-10-144-094-7

Query Match 66.1%; Score 21.8; DB 17; Length 33;
 Best Local Similarity 78.8%; Pred. No. 2;
 Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGTCGAAATAGTGAGT 33
 DB 1 CATCTCTCTCTCCGAGCCGTCGAAATAGTGAGT 33

```
RESULT 16
US-10-074-475-102/c
; Sequence 102, Application US/10074475
; Publication No. US2003002898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: US/10/074,475
; PRIOR FILING DATE: 2002-02-13
; NUMBER OF SEQ ID NOS: 295
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 102
; LENGTH: 698
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-102
Query Match 60.6%; Score 20; DB 15; Length 698;
Best Local Similarity 82.1%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCGCGTGGAAATAG 28
Db 34 CAGCTCTGCTCCCAAGCGCGTGGAAATGG 7

RESULT 17
US-10-369-493-28205/c
; Sequence 28205, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 28205
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Burkholderia fungorum
US-10-369-493-28205
Query Match 60.6%; Score 20; DB 16; Length 765;
Best Local Similarity 82.1%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCGCGTGGAAATAG 28
Db 84 CTTCGTCTCTCCGAGCGCGTGGAAATAG 57

RESULT 18
US-10-369-493-30964/c
; Sequence 30964, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 30964
; LENGTH: 771
; TYPE: DNA
; ORGANISM: Burkholderia cepacia
US-10-369-493-30964
Query Match 60.6%; Score 20; DB 16; Length 771;
Best Local Similarity 82.1%; Pred. No. 20;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATCTCTTCTCCGAGCGCGTGGAAATAG 28
Db 93 CTTCGTCTCTCCGAGCGCGTGGAAATAG 66

RESULT 19
US-09-974-300-1317/c
; Sequence 1317, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Id Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085-500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/683,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1317
; LENGTH: 1018
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1317
Query Match 60.6%; Score 20; DB 9; Length 1018;
Best Local Similarity 82.1%; Pred. No. 21;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 ATCTCTTCTCCGAGCGCGTGGAAATAGT 29
Db 845 ATGACTTCTCCGAGCGCGTGGAAATAGT 818

RESULT 20
US-10-074-475-103/c
; Sequence 103, Application US/10074475
; Publication No. US2003002898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
```

/ APPLICANT: Cafferkey, Robert
 / APPLICANT: Sun, Yongming
 / APPLICANT: Liu, Chenghua
 / TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
 / TITLE OF INVENTION: Genes and Proteins
 / FILE REFERENCE: DEX-0313
 / CURRENT APPLICATION NUMBER: US/10/074,475
 / PRIOR FILING DATE: 2002-02-13
 / PRIOR APPLICATION NUMBER: 60/266,292
 / NUMBER OF SEQ ID NOS: 295
 / SOFTWARE: PatentIn version 3.1
 / SEQ ID NO 103
 / LENGTH: 1217
 / TYPE: DNA
 / ORGANISM: Homo sapien
 US-10-074-475-103

Query Match 60.6%; Score 20; DB 15; Length 1217;
 Best Local Similarity 82.1%; Pred. No. 21;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CATCTCTCTCCGAGCCGCTCGAATAG 28
 Db 34 CAGCTCTCTCCGAGCCGCTCGAATAG 7

RESULT 21
 US-09-939-964-1
 / Sequence 1, Application US/09939964
 / Publication No. US20030054522A1
 / GENERAL INFORMATION:
 / APPLICANT: Rosenthal, Andre
 / APPLICANT: Freiberg, Christoph
 / APPLICANT: Perret, Xavier Philippe
 / APPLICANT: Broughton, William John
 / TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
 / TITLE OF INVENTION: Plasmid
 / FILE REFERENCE: CARP0068
 / CURRENT APPLICATION NUMBER: US/09/939,964
 / CURRENT FILING DATE: 2001-08-27
 / PRIOR APPLICATION NUMBER: 09/214,808
 / PRIOR FILING DATE: 1999-06-22
 / NUMBER OF SEQ ID NOS: 1
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 1
 / LENGTH: 536165
 / TYPE: DNA
 / ORGANISM: Rhizobium
 US-09-939-964-1

Query Match 60.6%; Score 20; DB 10; Length 536165;
 Best Local Similarity 82.1%; Pred. No. 42;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 CTCTCTCTCCGAGCCGCTCGAATAGTGA 31
 Db 455980 CTCTTCAGCGAGCCGCTCGACCTCTGTA 455907

RESULT 22
 US-10-424-599-50726
 / Sequence 50726, Application US/10424599
 / Publication No. US20040031072A1
 / GENERAL INFORMATION:
 / APPLICANT: La Rosa Thomas J
 / APPLICANT: Kovalic David K
 / APPLICANT: Zhou Yihua
 / APPLICANT: Cao Yongwei
 / TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / FILE REFERENCE: 38-21(5323)B
 / CURRENT APPLICATION NUMBER: US/10/424,599

/ CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 285684
 / SEQ ID NO 50726
 / LENGTH: 453
 / TYPE: DNA
 / ORGANISM: Glycine max
 / FEATURE:
 / OTHER INFORMATION: Clone ID: PAT_MKT3847_16816C.1
 US-10-424-599-50726

Query Match 60.0%; Score 19.8; DB 13; Length 453;
 Best Local Similarity 77.4%; Pred. No. 23;
 Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 2 ATCTCTCTCCGAGCCGCTCGAATAGTGA 32
 Db 5 ATCTCATCCCGAGCTCATCGAATCTCTGAG 35

RESULT 23
 US-10-425-114-30662
 / Sequence 30662, Application US/10425114
 / Publication No. US20040034888A1
 / GENERAL INFORMATION:
 / APPLICANT: Liu, Jinqdong
 / APPLICANT: Zhou, Yihua
 / APPLICANT: Kovalic, David K.
 / APPLICANT: Screen, Steven E.
 / APPLICANT: Tabashnik, Jack E.
 / APPLICANT: Cao, Yongwei
 / TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
 / TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 / FILE REFERENCE: 38-21(53313)B
 / CURRENT APPLICATION NUMBER: US/10/425,114
 / CURRENT FILING DATE: 2003-04-28
 / NUMBER OF SEQ ID NOS: 73128
 / SEQ ID NO 30662
 / LENGTH: 2581
 / TYPE: DNA
 / ORGANISM: Zea mays
 / FEATURE:
 / OTHER INFORMATION: Clone ID: UC-ZMPLB73059H03_FLI
 US-10-425-114-30662

Query Match 60.0%; Score 19.8; DB 13; Length 2581;
 Best Local Similarity 91.3%; Pred. No. 29;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATCTCTCTCCGAGCCGCTCGA 23
 Db 591 CAACTCTCTCCGAGCCGCTCTA 613

RESULT 24
 US-09-782-378A-27/c
 / Sequence 27, Application US/09782378A
 / Patent No. US20020102731A1
 / GENERAL INFORMATION:
 / APPLICANT: Hearing, Patrick
 / APPLICANT: Bahou, Nadie
 / APPLICANT: Sandalon, Ziv
 / APPLICANT: Gnatensko, Dmitri
 / TITLE OF INVENTION: Adenoviral Vectors
 / FILE REFERENCE: STONY-04970
 / CURRENT APPLICATION NUMBER: US/09/782,378A
 / CURRENT FILING DATE: 2001-02-12
 / PRIOR APPLICATION NUMBER: 60/237,747
 / PRIOR FILING DATE: 2000-10-02
 / NUMBER OF SEQ ID NOS: 27
 / SOFTWARE: PatentIn version 3.0
 / SEQ ID NO 27
 / LENGTH: 34214
 / TYPE: DNA

ORGANISM: Human adenovirus type 40
US-09-782-378A-27

Query Match 58.8%; Score 19.4; DB 9; Length 34214;
Best Local Similarity 79.3%; Pred. No. 59;
Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATCTCTTCGAGCGGTCGAATAGTG 30
Db 12904 ATTTTTCGCGCCGTCACATAGG 12876

RESULT 25

US-10-144-679-14
; Sequence 14, Application US/10144679
; Publication No. US20030215810A1
; GENERAL INFORMATION:
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
; TITLE OF INVENTION: COLOR CHANGES
; FILE REFERENCE: 9800240-0019
; CURRENT APPLICATION NUMBER: US/10/144,679
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-144-679-14

Query Match 58.2%; Score 19.2; DB 16; Length 50;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ATCTCTTCGAGCGGTCGAATAGTG 33
Db 1 ATCTCTTTGTGACCGACTCGAATAGTG 32

RESULT 26

US-10-144-094-14
; Sequence 14, Application US/10144094
; Publication No. US20040023216A1
; GENERAL INFORMATION:
; APPLICANT: LIU, JUEWEN
; TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
; FILE REFERENCE: 10322/44
; CURRENT APPLICATION NUMBER: US/10/144,094
; CURRENT FILING DATE: 2002-05-10
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 50
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA
US-10-144-094-14

Query Match 58.2%; Score 19.2; DB 17; Length 50;
Best Local Similarity 75.0%; Pred. No. 35;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2 ATCTCTTCGAGCGGTCGAATAGTG 33
Db 1 ATCTCTTTGTGACCGACTCGAATAGTG 32

RESULT 27

US-10-156-761-1/c
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIIKAWA, JUN
; APPLICANT: HOKIYAMA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-156-761-1

Query Match 58.2%; Score 19.2; DB 15; Length 9025608;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CATCTCTTCGAGCGGTCGAATAGTG 32
Db 8349509 CCTCGCTTACCGCTGCTGTCGAATAGTG 8349478

RESULT 28

US-09-796-692-4311/c
; Sequence 4311, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mansion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/260,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/260,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/262,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903

;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 4311
;; LENGTH: 444
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (31)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (34)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (36)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (46)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (64)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (69)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (71)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (100)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (115)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (121)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (164)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (178)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (184)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (192)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (213)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (221)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (251)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (278)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (292)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure

;; LOCATION: (299)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (315)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (317)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (385)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (413)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (417)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (428)
;; OTHER INFORMATION: n=A,T,C or G
;; US-09-796-692-4311

Query Match 57.6%; Score 19; DB 9; Length 444;
Best Local Similarity 78.6%; Pred. No. 56;
Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4 CTCCTCTCCGAGCGGTCGAATAGTGA 31
||| ||||| ||||| ||||| ||||| |||||
DB 151 CTCCTCTCCGAGCGGTCGAATAGTGA 124

RESULT 29
US-10-640-862-4311/c
;; Sequence 4311, Application US/10040862
;; Publication No. US20030078396A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Retter, Marc
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;; FILE REFERENCE: 014058-013520US
;; CURRENT APPLICATION NUMBER: US/10/040,862
;; CURRENT FILING DATE: 2001-11-06
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: US 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: US 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; PRIOR APPLICATION NUMBER: US 09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; NUMBER OF SEQ ID NOS: 10467

```

1 SOFTWARE: FastSeq for Windows Version 3.0
2 SEQ ID NO 4311
3 LENGTH: 444
4
5 TYPE: DNA
6
7 ORGANISM: Homo sapiens
8
9 FEATURE:
10 NAME/KEY: unsure
11 LOCATION: (31)
12 OTHER INFORMATION: n=A,T,C or G
13
14 FEATURE:
15 NAME/KEY: unsure
16 LOCATION: (34)
17 OTHER INFORMATION: n=A,T,C or G
18
19 FEATURE:
20 NAME/KEY: unsure
21 LOCATION: (36)
22 OTHER INFORMATION: n=A,T,C or G
23
24 FEATURE:
25 NAME/KEY: unsure
26 LOCATION: (46)
27 OTHER INFORMATION: n=A,T,C or G
28
29 FEATURE:
30 NAME/KEY: unsure
31 LOCATION: (64)
32 OTHER INFORMATION: n=A,T,C or G
33
34 FEATURE:
35 NAME/KEY: unsure
36 LOCATION: (69)
37 OTHER INFORMATION: n=A,T,C or G
38
39 FEATURE:
40 NAME/KEY: unsure
41 LOCATION: (71)
42 OTHER INFORMATION: n=A,T,C or G
43
44 FEATURE:
45 NAME/KEY: unsure
46 LOCATION: (100)
47 OTHER INFORMATION: n=A,T,C or G
48
49 FEATURE:
50 NAME/KEY: unsure
51 LOCATION: (115)
52 OTHER INFORMATION: n=A,T,C or G
53
54 FEATURE:
55 NAME/KEY: unsure
56 LOCATION: (121)
57 OTHER INFORMATION: n=A,T,C or G
58
59 FEATURE:
60 NAME/KEY: unsure
61 LOCATION: (146)
62 OTHER INFORMATION: n=A,T,C or G
63
64 FEATURE:
65 NAME/KEY: unsure
66 LOCATION: (164)
67 OTHER INFORMATION: n=A,T,C or G
68
69 FEATURE:
70 NAME/KEY: unsure
71 LOCATION: (178)
72 OTHER INFORMATION: n=A,T,C or G
73
74 FEATURE:
75 NAME/KEY: unsure
76 LOCATION: (184)
77 OTHER INFORMATION: n=A,T,C or G
78
79 FEATURE:
80 NAME/KEY: unsure
81 LOCATION: (192)
82 OTHER INFORMATION: n=A,T,C or G
83
84 FEATURE:
85 NAME/KEY: unsure
86 LOCATION: (213)
87 OTHER INFORMATION: n=A,T,C or G
88
89 FEATURE:
90 NAME/KEY: unsure
91 LOCATION: (221)
92 OTHER INFORMATION: n=A,T,C or G

```

```

1 FEATURE:
2 NAME/KEY: unsure
3 LOCATION: [251]
4 OTHER INFORMATION: n=A,T,C or G
5 FEATURE:
6 NAME/KEY: unsure
7 LOCATION: [278]
8 OTHER INFORMATION: n=A,T,C or G
9 FEATURE:
10 NAME/KEY: unsure
11 LOCATION: [292]
12 OTHER INFORMATION: n=A,T,C or G
13 FEATURE:
14 NAME/KEY: unsure
15 LOCATION: [299]
16 OTHER INFORMATION: n=A,T,C or G
17 FEATURE:
18 NAME/KEY: unsure
19 LOCATION: [315]
20 OTHER INFORMATION: n=A,T,C or G
21 FEATURE:
22 NAME/KEY: unsure
23 LOCATION: [317]
24 OTHER INFORMATION: n=A,T,C or G
25 FEATURE:
26 NAME/KEY: unsure
27 LOCATION: [385]
28 OTHER INFORMATION: n=A,T,C or G
29 FEATURE:
30 NAME/KEY: unsure
31 LOCATION: [413]
32 OTHER INFORMATION: n=A,T,C or G
33 FEATURE:
34 NAME/KEY: unsure
35 LOCATION: [417]
36 OTHER INFORMATION: n=A,T,C or G
37 FEATURE:
38 NAME/KEY: unsure
39 LOCATION: [428]
40 OTHER INFORMATION: n=A,T,C or G
41 US-10-040-862-4311

```

Query Match	57.6%	Score 19;	DB 15;	Length 444;
Best Local Similarity	78.6%	Pred. No. 56;		
Matches 22;	Conservative	0;	Mismatches	6; Indels

OY
4 CTCTTCTCCGAGCCGGTCGAATAGTGA 31

D6
151 CTCCTNCCCGAGTAGGTAGAGATAGTGA 124

RESULT 30
US-10-057-475B-4311/c
/ Sequence 4311, Application US/10057475B
/ Publication No. US2004002068A1
/ CENTRAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Clapper, Jonathan David
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordones, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014056-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479

OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA
US-10-144-094-17

Query Match 57.0%, Score 18.8, DB 17, Length 51,
Best Local Similarity 76.7%, Pred. No. 55,
Matches 23, Conservative 0, Mismatches 7, Indels 0, Gaps 0;

QY 2 ATCTCTCTCCGAGCGGTGGAATAGTGA 31
DB 6 ATTACTTCTACGAGCGGTGGAATAGTGA 35

RESULT 34

US-10-424-599-74527
Sequence 74527, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 74527
LENGTH: 220
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_38311C.1
US-10-424-599-74527

Query Match 57.0%, Score 18.8, DB 13, Length 220;
Best Local Similarity 76.7%, Pred. No. 65,
Matches 23, Conservative 0, Mismatches 7, Indels 0, Gaps 0;

QY 1 CACTCTCTCCGAGCGGTGGAATAGTG 30
DB 153 CTTCACTTTCCGACGCGTTCGAACAGTG 182

RESULT 35

US-09-974-300-5654/c
Sequence 5654, Application US/03974300
Patent No. US20020146721A1
GENERAL INFORMATION:
APPLICANT: Betka, Randy M.
APPLICANT: Clausen, Ib Groth
TITLE OF INVENTION: Methods For Monitoring Multiple Gene
FILE OF INVENTION: Expression
FILE REFERENCE: 10085.500-US
CURRENT APPLICATION NUMBER: US/09/974,300
CURRENT FILING DATE: 2001-10-05
PRIOR APPLICATION NUMBER: 03/680,598
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/279,526
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 8481
SOFTWARE: PASTSEQ for Windows Version 4.0
SEQ ID NO 5654
LENGTH: 495
TYPE: DNA
ORGANISM: Bacillus clausii
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(495)
OTHER INFORMATION: n = A,T,C or G
US-09-974-300-5654

Query Match 57.0%, Score 18.8, DB 9, Length 495;

Best Local Similarity 76.7%, Pred. No. 71;
Matches 23, Conservative 0, Mismatches 7, Indels 0, Gaps 0;

QY 3 TCTCTCTCCGAGCGGTGGAATAGTGA 32
DB 120 TCTCTCTCCGAGCGGTGGAATAGTGA 91

RESULT 36

US-10-282-122A-14316
Sequence 14316, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

FILE REFERENCE: ELITRA-034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14316
LENGTH: 846
TYPE: DNA
ORGANISM: Burkholderia mallei
US-10-282-122A-14316

Query Match 57.0%, Score 18.8, DB 13, Length 846;
Best Local Similarity 76.7%, Pred. No. 75,
Matches 23, Conservative 0, Mismatches 7, Indels 0, Gaps 0;

QY 2 ATCTCTCTCCGAGCGGTGGAATAGTGA 31
DB 761 ATCGTACGACGCGGTGGAATAGTGA 790

RESULT 37

US-10-425-114-25452/c
Sequence 25452, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong

```

; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 25452
; LENGTH: 1132
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3910-027-D8_FLI
US-10-425-114-25452

Query Match 57.0%; Score 18.8; DB 13; Length 1132;
Best Local Similarity 76.7%; Pred. No. 78;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TCTCTTCCTCCGAGCCGCTCGAATAGTGAG 32
|||||
DB 1001 TCTCTCCACCGAGCCGCTAGAAAGTGG 972

RESULT 38
US-10-425-114-22309/c
; Sequence 22309, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven B
; APPLICANT: Tabaska, Jack B
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22309
; LENGTH: 1229
; TYPE: DNA
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3399-014-A6_FLI
US-10-425-114-22309

Query Match 57.0%; Score 18.8; DB 13; Length 1229;
Best Local Similarity 76.7%; Pred. No. 79;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 3 TCTCTTCCTCCGAGCCGCTCGAATAGTGAG 32
|||||
DB 1098 TCTCTCCACCGAGCCGCTAGAAAGTGG 1069

RESULT 39
US-10-282-122A-14254
; Sequence 14254, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangueu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith

```

```

; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EPIGRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14254
; LENGTH: 1371
; TYPE: DNA
; ORGANISM: Burkholderia mallei
US-10-282-122A-14254

Query Match 57.0%; Score 18.8; DB 13; Length 1371;
Best Local Similarity 90.9%; Pred. No. 80;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCTCTTCCTCCGAGCCGCTCGA 23
|||||
DB 380 ATCTCTTCCTCCGAGCCGCTCGA 401

RESULT 40
US-10-194-163-1029/c
; Sequence 1029, Application US/10194163
; Publication No. US20020172976A1
; GENERAL INFORMATION:
; APPLICANT: Rose, Bruce Carter
; TITLE OF INVENTION: FOREHYMONAS GINGIVALIS ECLYNUCLEOTIDES
; AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSES: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: PALO ALTO
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/194,163
; FILING DATE: 04-Nov-2002
; CLASSIFICATION: <Unknown>

```

```

/ ATTORNEY/AGENT INFORMATION:
/ NAME: Basu, Shantanu
/ REGISTRATION NUMBER: 43,318
/ REFERENCE/DOCKET NUMBER: 529282000101
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-813-5995
/ TELEFAX: 650-494-0792
/ TELEX: 706141
/ INFORMATION FOR SEQ ID NO: 1029
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 3182 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: circular
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: UNKNOWN
/ ORIGINAL SOURCE:
/ ORGANISM: PORPHYROMONAS GINGIVALIS
/ FEATURES:
/ NAME/KEY: misc feature
/ LOCATION: 1..3182
/ SEQUENCE DESCRIPTION: SEQ ID NO: 1029
US-10-144-163-1029

```

```

Query Match      57.04; Score 18.8; DB 13; Length 3182;
Best Local Similarity 76.74; Pred. NO. 87;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Ov 1 CATCTCTTCCTCCGAGCCGTCGAAATAGTG 30
   |||||
Db 1334 CAGCTCTTCTCCGATCGCTGATATAGTG 1305

```

```

Search completed: May 24, 2004, 13:45:29
Job time : 216.245 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 07:43:51 ; Search time 202.981 seconds
(without alignments)
690.658 Million cell updates/sec

Title: US-10-144-679-1

Perfect score: 33
Sequence: 1 catctcttcgagccggtcgaaatagtgagt 33

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_25Jan04 :
1: Geneseq1980s :
2: Geneseq1990s :
3: Geneseq2000s :
4: Geneseq2001as :
5: Geneseq2001bs :
6: Geneseq2002s :
7: Geneseq2003as :
8: Geneseq2003bs :
9: Geneseq2003cs :
10: Geneseq2004s :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	100.0	33	6 AAL45459	Aal45459 RNA-cleav
2	33	100.0	33	6 AAL45380	Aal45380 Zn(II)-de
3	27.8	84.2	50	6 AAL45398	Aal45398 Zinc-depe
4	26.4	80.0	50	6 AAL45397	Aal45397 Zinc-depe
5	26.4	80.0	50	6 AAL45396	Aal45396 Zinc-depe
6	24	72.7	50	6 AAL45399	Aal45399 Zinc-depe
7	21.8	66.1	33	6 AAL45386	Aal45386 Zn(II)-de
8	20.6	62.4	475	4 AAL45388	Aal45388 Human GAB
9	20.6	62.4	1248	6 AAL47338	Aal47338 Human GAB
10	20.6	62.4	2616	2 AAL47338	Aal47338 Human GAB
11	20.4	61.8	1254	9 AAL47338	Aal47338 Human GAB
12	20.2	61.2	2695	9 AAL47338	Aal47338 Human GAB
13	20.2	61.2	2695	9 AAL47338	Aal47338 Human GAB
14	20.6	60.6	698	6 AAL47338	Aal47338 Human GAB
15	20.6	60.6	1018	6 AAL47338	Aal47338 Human GAB
16	20.6	60.6	1217	6 AAL47338	Aal47338 Human GAB
17	20.6	60.6	110000	2 AAL47338	Aal47338 Human GAB
18	19.8	60.0	1230	3 AAL47338	Aal47338 Human GAB
19	19.4	58.8	34214	6 AAL47338	Aal47338 Human GAB
20	19.2	58.2	50	6 AAL45392	Aal45392 Zinc-depe
21	19.5	57.6	1512	4 AAL45392	Aal45392 Zinc-depe
22	19.5	57.6	3686	4 AAL45392	Aal45392 Zinc-depe
23	18.8	57.0	50	2 AAL45392	Aal45392 Zinc-depe

24	18.8	57.0	51	6 AAL45395	Aal45395 Zinc-depe
25	18.8	57.0	495	6 AAL45395	Aal45395 Zinc-depe
26	18.8	57.0	636	2 AAL45395	Aal45395 Zinc-depe
27	18.8	57.0	734	8 AAL45395	Aal45395 Zinc-depe
28	18.8	57.0	734	9 AAL45395	Aal45395 Zinc-depe
29	18.8	57.0	846	7 AAL45395	Aal45395 Zinc-depe
30	18.8	57.0	1371	7 AAL45395	Aal45395 Zinc-depe
31	18.8	57.0	1926	6 AAL45395	Aal45395 Zinc-depe
32	18.6	56.4	49	6 AAL45395	Aal45395 Zinc-depe
33	18.6	56.4	49	6 AAL45395	Aal45395 Zinc-depe
34	18.6	56.4	54	6 AAL45395	Aal45395 Zinc-depe
35	18.6	56.4	54	6 AAL45395	Aal45395 Zinc-depe
36	18.6	56.4	512	9 AAL45395	Aal45395 Zinc-depe
37	18.6	56.4	1024	6 AAL45395	Aal45395 Zinc-depe
38	18.6	56.4	1024	6 AAL45395	Aal45395 Zinc-depe
39	18.6	56.4	1024	6 AAL45395	Aal45395 Zinc-depe
40	18.6	56.4	1035	2 AAL45395	Aal45395 Zinc-depe
41	18.6	56.4	1311	6 AAL45395	Aal45395 Zinc-depe
42	18.6	56.4	1626	6 AAL45395	Aal45395 Zinc-depe
43	18.6	56.4	1626	6 AAL45395	Aal45395 Zinc-depe
44	18.6	56.4	1722	3 AAL45395	Aal45395 Zinc-depe
45	18.6	56.4	1722	3 AAL45395	Aal45395 Zinc-depe

ALIGNMENTS

RESULT 1
AAL45459
ID AAL45459 standard; DNA; 33 BP.
XX
AC AAL45459;
XX
DT 06-JUN-2002 (first entry)
XX
DE RNA-cleaving deoxyribozyme #7.
XX
XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
XX quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.
XX Unidentified.
XX
XX Key Location/Qualifiers
XX misc_binding 1..9
XX /tag= a
XX /bound_moiety= "substrate"
XX /note= "binds nucleotides 20-12 of the substrate shown in
XX AAL45460"
XX stem_loop 11..19
XX /tag= b
XX misc_binding 25..33
XX /tag= c
XX /bound_moiety= "substrate"
XX /note= "binds nucleotides 9-1 of the substrate shown in
XX AAL45460".
XX
XX W0200200006-A2.
XX
XX 03-JAN-2002.
XX
XX 27-JUN-2001; 2001WO-US020557.
XX
XX 27-JUN-2000; 2000US-00605558.
XX
XX (UNII) UNIV ILLINOIS FOUND.
XX
XX Lu Y, Li J;
XX WPI; 2002-130823/17.
XX
XX New nucleic acid enzyme biosensors, useful for the sensitive and
XX selective detection of ions, particularly metal ions e.g. lead ions, and
XX for determining the concentration of a particular ion in a solution.


```

XX PS      Example 1; Fig 6; 57pp; English.
XX CC      The present invention relates to biosensors, comprising a nucleic acid
XX CC      enzyme dependent on an ion to produce a product, a quencher or/and a
XX CC      fluorophore and a photodetector. The biosensors are useful for the
XX CC      sensitive and selective detection of ions. The biosensors are useful in
XX CC      methods of detecting the presence of an ion, particularly metal ions such
XX CC      as lead. The biosensors may also be used to determine the concentration
XX CC      of a particular ion in a solution. The present sequence is a ribozyme
XX CC      which may be used as a biosensor of the invention
XX SQ      Sequence 33 BP; 7 A; 9 C; 8 G; 9 T; 0 U; 0 Other;
XX
XX          Query Match      100.0%; Score 33; DB 6; Length 33;
XX          Best Local Similarity 100.0%; Pred. No. 4.3e-05;
XX          Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY      1 CATCTCTCTCCGAGCCGCTCCGAAATAGTGAGT 33
XX DB      1 CATCTCTCTCCGAGCCGCTCCGAAATAGTGAGT 33
XX
XX RESULT 2
XX AAL45380
XX ID      AAL45380 standard; DNA; 33 BP.
XX AC      AAL45380;
XX XX
XX DT      06-JUN-2002 (first entry)
XX DE      Zn(II)-dependent trans-cleaving deoxyribozyme substrate 17DS.
XX KW      Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
XX KW      quencher; fluorophore; photodetector; ion concentration; ribozyme;
XX KW      substrate; ss.
XX OS      Unidentified.
XX
XX PH      Key
XX FT      misc_binding
XX FT      Location/Qualifiers
XX FT      1..9
XX FT      /tag= a
XX FT      /bound_moiety= "substrate 17DS"
XX FT      /note= "binds nucleotides 20-12 of substrate 17DS shown
XX FT      in AAL45381"
XX FT      10..18
XX FT      /tag= b
XX FT      misc_binding
XX FT      24..32
XX FT      /tag= c
XX FT      /bound_moiety= "substrate 17DS"
XX FT      /note= "binds nucleotides 9-1 of substrate 17DS shown in
XX FT      AAL45381"
XX PN      WO200200006-A2.
XX XX
XX PD      03-JAN-2002.
XX PF      27-JUN-2001; 2001WO-US020557.
XX PR      27-JUN-2000; 2000US-00605558.
XX XX
XX PA      (UNII ) UNIV ILLINOIS FOUND.
XX PI      Lu Y, Li J;
XX DR      WPI; 2002-130823/17.
XX
XX PT      New nucleic acid enzyme biosensors, useful for the sensitive and
XX PT      selective detection of ions, particularly metal ions e.g. lead ions, and
XX PT      for determining the concentration of a particular ion in a solution.
XX PS      Claim 17; Fig 5; 57pp; English.
XX

```

```

CC CC      The present invention relates to biosensors, comprising a nucleic acid
CC CC      enzyme dependent on an ion to produce a product, a quencher or/and a
CC CC      fluorophore and a photodetector. The biosensors are useful for the
CC CC      sensitive and selective detection of ions. The biosensors are useful in
CC CC      methods of detecting the presence of an ion, particularly metal ions such
CC CC      as lead. The biosensors may also be used to determine the concentration
CC CC      of a particular ion in a solution. The present sequence is a ribozyme
CC CC      which may be used as a biosensor of the invention
XX SQ      Sequence 33 BP; 7 A; 9 C; 8 G; 9 T; 0 U; 0 Other;
XX
XX          Query Match      100.0%; Score 33; DB 6; Length 33;
XX          Best Local Similarity 100.0%; Pred. No. 4.3e-05;
XX          Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY      1 CATCTCTCTCCGAGCCGCTCCGAAATAGTGAGT 33
XX DB      1 CATCTCTCTCCGAGCCGCTCCGAAATAGTGAGT 33
XX
XX RESULT 3
XX AAL45398
XX ID      AAL45398 standard; DNA; 50 BP.
XX AC      AAL45398;
XX XX
XX DT      06-JUN-2002 (first entry)
XX DE      Zinc-dependent deoxyribozyme SEQ ID NO: 20.
XX KW      Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
XX KW      quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.
XX OS      Unidentified.
XX PN      WO200200006-A2.
XX PD      03-JAN-2002.
XX PF      27-JUN-2001; 2001WO-US020557.
XX PR      27-JUN-2000; 2000US-00605558.
XX PA      (UNII ) UNIV ILLINOIS FOUND.
XX PI      Lu Y, Li J;
XX DR      WPI; 2002-130823/17.
XX
XX PT      New nucleic acid enzyme biosensors, useful for the sensitive and
XX PT      selective detection of ions, particularly metal ions e.g. lead ions, and
XX PT      for determining the concentration of a particular ion in a solution.
XX XX
XX Example 1; Fig 2; 57pp; English.
XX
XX CC      The present invention relates to biosensors, comprising a nucleic acid
XX CC      enzyme dependent on an ion to produce a product, a quencher or/and a
XX CC      fluorophore and a photodetector. The biosensors are useful for the
XX CC      sensitive and selective detection of ions. The biosensors are useful in
XX CC      methods of detecting the presence of an ion, particularly metal ions such
XX CC      as lead. The biosensors may also be used to determine the concentration
XX CC      of a particular ion in a solution. The present sequence is a ribozyme
XX CC      which may be used as a biosensor of the invention
XX SQ      Sequence 50 BP; 9 A; 12 C; 14 G; 15 T; 0 U; 0 Other;
XX
XX          Query Match      84.2%; Score 27.8; DB 6; Length 50;
XX          Best Local Similarity 93.5%; Pred. No. 0.01;
XX          Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY      2 ATCTCTCTCCGAGCCGCTCCGAAATAGTGAG 32
XX DB      1 ATCTCTCTCCGAGCCGCTCCGAAATAGTGAG 31
XX

```


XX The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependant on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a ribozyme
 CC which may be used as a biosensor of the invention

XX Sequence 50 BP; 12 A; 11 C; 10 G; 17 T; 0 U; 0 Other;

Query Match 72.7%; Score 24; DB 6; Length 50;

Best Local Similarity 100.0%; Pred. No. 0.53;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 CTCTCTCCGAGCCGGTCGAAATAGT 29

DB 6 CTCTCTCCGAGCCGGTCGAAATAGT 29

RESULT 7

AA45386

ID AAL45386 standard; DNA; 33 BP.

XX AAL45386;

XX 06-JUN-2002 (first entry)

XX Zn(II)-dependent trans-cleaving deoxyribozyme 1781.

XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;

XX quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.

XX Unidentified.

XX WO200200006-A2.

XX 03-JAN-2002.

XX 27-JUN-2001; 2001WO-US020557.

XX 27-JUN-2000; 2000US-00605558.

XX (UNII) UNIV ILLINOIS FOUND.

XX Lu Y, Li J;

XX WPI; 2002-130823/17.

XX New nucleic acid enzyme biosensors, useful for the sensitive and
 CC selective detection of ions, particularly metal ions e.g. lead ions, and
 CC for determining the concentration of a particular ion in a solution.

XX Example 1; Page 24; 57pp; English.

XX The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependant on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a ribozyme
 CC which may be used as a biosensor of the invention

XX Sequence 33 BP; 8 A; 7 C; 7 G; 11 T; 0 U; 0 Other;

Query Match 66.1%; Score 21.8; DB 6; Length 33;

Best Local Similarity 78.8%; Pred. No. 4.9;

Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGT 33

|||||

DB 1 CATCTCTCTCTCCGAGCCGGTCGAAATAGT 33

RESULT 8

AAH88480

ID AAH88480 standard; DNA; 475 BP.

XX AAH88480;

XX 26-FEB-2002 (first entry)

XX CNS disorder-related biallelic marker #2 from GABRG2 gene.

XX Single nucleotide polymorphism; SNP; biallelic marker; human; ds;

XX central nervous system disorder; CNS; GABRG2; GABA-A receptor subunit.

XX Homo sapiens.

XX Key Location/Qualifiers

XX misc_feature 352

XX /tag= a

XX /standard_name= "single nucleotide polymorphism"

XX WO200151659-A2.

XX 19-JUL-2001.

XX 11-JAN-2001; 2001WO-18000116.

XX 13-JAN-2000; 2000US-0175854P.

XX (GIST) GENSET.

XX Chu T, Blumanfeld M, Cohen D;

XX WPI; 2001-483085/52.

XX Isolated polymucleotides, useful for genotyping nucleic acids for
 CC biallelic markers for the diagnosis of depression, comprises central
 CC nervous system disorder related biallelic marker.

XX Claim 1; Page 347; 519pp; English.

XX The present invention relates to biallelic markers derived from human
 CC genes involved in central nervous system (CNS) disorders. The present
 CC sequence is one such biallelic marker derived from human GABA-A receptor
 CC subunit (GABRG2) gene. GABA is gamma aminobutyric acid. This marker has a
 CC single nucleotide polymorphism (SNP) and is useful in determining the
 CC genetic predisposition of individuals to CNS disorders by identifying
 CC the nucleotides at a set of genetic markers in a biological sample, where
 CC the markers comprise at least one CNS disorder related marker

XX Sequence 475 BP; 137 A; 92 C; 98 G; 146 T; 0 U; 2 Other;

Query Match 62.4%; Score 20.6; DB 4; Length 475;

Best Local Similarity 85.2%; Pred. No. 26;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TCCTCTCTCCGAGCCGGTCGAAATAGT 29

DB 11 TCCTCTCTCCGAGCCGGTCGAAATAGT 37

RESULT 9

AA47338

ID AAL47338 standard; DNA; 1248 BP.

XX AC AAL47338;

XX 05-SEP-2002 (first entry)

XX Human GABP beta-HSV1 VP16 fusion protein coding sequence.

XX

KW Human; utrophin B, promoter; Ets-transcription factor; transactivator;
 KW muscular disease; Duchenne muscular dystrophy; Becker muscular dystrophy;
 KW gene therapy; muscle wasting; cytoskeletal; immunostimulant; haemostatic;
 KW haemophilia; immune deficiency; cancer; HSV1; VP16; GABP beta; ds;
 KW fusion protein.

XX Homo sapiens.
 CS Herpes simplex virus; type 1.
 OS Synthetic.
 OS Chimeric.

PH Key Location/Qualifiers
 FT CDS 1..1248
 FT /tag= a
 FT /product= "GABPbeta_VP16 fusion protein"

XX WO200236620-A2.

XX 10-MAY-2002.

XX 31-OCT-2001; 2001WO-EP012662.

XX 02-NOV-2000; 2000EP-00123842.

XX (MYOC-) MYOCONTRACT PHARM RES AG.

XX Rueegg MA, Briguet A;

XX WPI; 2002-500123/53.

XX P-PSDB; AAO18052.

XX Novel nucleic acid capable of controlling the expression of a gene and
 PT being activated by an Ets-transcription factor related compound, useful
 PT for screening candidate compounds capable of regulating transcription.

XX Disclosure; Fig 2D; 38pp; English.

XX The present invention provides a nucleic acid capable of controlling the
 CC expression of a gene and being activated by an Ets-transcription factor
 CC (Etr) related compound. The nucleic acid is useful for screening and/or
 CC providing candidate compounds capable of regulating transcription. This
 CC involves bringing the nucleic acid into contact with compounds to be
 CC screened and detecting the transcriptional activity from the nucleic acid
 CC in the presence and absence of the compounds and optionally purifying
 CC and/or synthesizing the positively tested compound. A construct
 CC containing the sequence is useful for activating the expression of a gene
 CC under its control, such as the utrophin gene, interleukin (IL)-2 gene,
 CC factor IX gene, CD18 gene, thrombopoietin (TPO) gene, Fas gene or
 CC acetylcholine receptor (AChR) delta and epsilon subunits genes, for the
 CC treatment of muscle diseases, preferably Duchenne or Becker muscular
 CC dystrophies, haemophilia, immune deficiency and cancer. The present
 CC sequence is the coding sequence of a fusion protein of the human GABP
 CC beta protein and Herpes simplex virus type 1 VP16

XX Sequence 1248 BP; 336 A; 285 C; 340 G; 287 T; 0 U; 0 Other;

Query Match 62.4%; Score 20.6; DB 6; Length 1248;
 Best Local Similarity 85.2%; Pred. No. 30;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 5 TCTTCTCCGAGCCGGTGGAAATAGTGA 31
 DB 168 TCTTCTCCGAGCCGGTGGAAATAGTGA 194

RESULT 10

AAQ37966

ID AAQ37966 standard; cDNA; 2615 BP.

XX AAQ37966;

XX 25-MAR-2003 (revised)

DT 30-JUN-1993 (first entry)

XX Sequence encoding purine-rich repeat (GA repeat) binding protein (GABP)
 DE subunit beta 1.
 DE GA binding protein; cis-regulatory element; VP16 mediated induction; ss.
 KW Mus musculus.

XX Key Location/Qualifiers
 FT CDS 133..1281
 FT /tag= a

XX WO9304166-A1.

XX 04-MAR-1993.

XX 17-AUG-1992; 92WO-US006748.

XX 16-AUG-1991; 91US-00746032.

XX (CARN-) CARNEGIE INST WASHINGTON.

XX McKnight SL, Thompson CC, Lamarco KL;

XX WPI; 1993-093998/11.

XX P-PSDB; AAR33366.

XX DNA encoding GA binding protein sub-unit - allows investigation of sub-
 PT unit sequence motif functions, for control of rapid cell division e.g. in
 PT cancer.

XX Claim 7; Fig 2B1-3; 68pp; English.

XX A cis-regulatory element required for virion associated protein VP16
 CC mediated induction of herpes simplex virus 1 (HSV1) immediate early (IE)
 CC genes consists of three imperfect repeats of the purine-rich
 CC hexanucleotide 5'-CGAAR-3'. A protein complex capable of avid
 CC interaction with the purine-rich repeats (GA repeats) has been identified
 CC in soluble preparations of rat liver nuclei. This GA binding protein
 CC (GABP) consists of two separable subunits. Applicants have isolated cDNA
 CC clones encoding both subunits of GABP and have revealed that one (GABP
 CC alpha) is related to the Ets transforming protein, while the other (GABP
 CC beta) contains a series of 31-amino acid repeats related in sequence to a
 CC variety of proteins. The sequences for GABP-beta-1 and GABP-beta-2 are
 CC identical up to nucleotide 1130 except for a three nucleotide insertion
 CC (GTA) at posn. 828 of GABP-beta-1. (Updated on 25-MAR-2003 to correct PN
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)

XX Sequence 2616 BP; 742 A; 557 C; 593 G; 724 T; 0 U; 0 Other;

Query Match 62.4%; Score 20.6; DB 2; Length 2616;
 Best Local Similarity 85.2%; Pred. No. 34;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 5 TCTTCTCCGAGCCGGTGGAAATAGTGA 31
 DB 300 TCTTCTCCGAGCCGGTGGAAATAGTGA 326

RESULT 11

ADC03463

ID ADC03463 standard; DNA; 1254 BP.

XX ADC03463;

XX 18-DEC-2003 (first entry)

XX Wheat flowering time-related DNA sequence #23.

XX flowering time; wheat; flower architecture; plant biotechnology;
 KW commercial plant farming; agriculture; flowering-related protein; ds.

XX Triticum aestivum.

XX WO2003000904-A2.
 XX 03-JAN-2003.
 XX 24-JUN-2002; 2002WO-EP006968.
 XX 22-JUN-2001; 2001US-0300112P.
 XX 26-SEP-2001; 2001US-0325277P.
 XX 30-NOV-2001; 2001US-0334984P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;
 XX Glazebrook J, Katagiri F, Krepes J, Provart N, Rieke D, Zhu T;
 XX WPI; 2003-229340/22.
 XX New isolated nucleic acid molecule encoding a polypeptide modulating
 XX flower architecture and flowering time, useful in the area of plant
 XX biotechnology, and commercial plant farming and agriculture.
 XX Disclosure; SEQ ID NO 103; 323pp; English.
 XX The invention comprises the amino acid and coding sequences of proteins
 XX involved in the control of flowering time in rice. The DNA and protein
 XX sequences of the invention are useful for modulating flower architecture
 XX and flowering time, the DNA and protein sequences are useful in the area
 XX of plant biotechnology, commercial plant farming and agriculture. The
 XX present wheat DNA sequence shows homology to the rice sequences of the
 XX invention.
 XX Sequence 1254 BP; 301 A; 296 C; 335 G; 322 T; 0 U; 0 Other;
 XX Query Match 61.8%; Score 20.4; DB 9; Length 1254;
 XX Best Local Similarity 80.0%; Pred. No. 37;
 XX Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGTG 30
 DB 738 CTCTCTCTCTCCGAGCCGGTCGAAATAGTG 767
 RESULT 12
 ID ADD48766/c
 XX ADD48766 standard; DNA; 2695 BP.
 XX AC ADD48766;
 XX 23-JAN-2004 (first entry)
 XX Human gene S75037; SEQ ID NO 14476.
 XX Human; ds; gene; pain; neuronal tissue; gene therapy;
 XX spinal segmental nerve injury; chronic constriction injury; CCI;
 XX spared nerve injury; SNI; Chung.
 XX Homo sapiens.
 XX WO2003016475-A2.
 XX 27-FEB-2003.
 XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 XX 01-NOV-2001; 2001US-0346382P.
 XX 26-NOV-2001; 2001US-0333347P.
 XX (GEHO) GEN HOSPITAL CORP.
 XX (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 PI WO2003000904-A2.
 XX 03-JAN-2003.
 XX 24-JUN-2002; 2002WO-EP006968.
 XX 22-JUN-2001; 2001US-0300112P.
 XX 26-SEP-2001; 2001US-0325277P.
 XX 30-NOV-2001; 2001US-0334984P.
 XX (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T;
 XX Glazebrook J, Katagiri F, Krepes J, Provart N, Rieke D, Zhu T;
 XX WPI; 2003-229340/22.
 XX New isolated nucleic acid molecule encoding a polypeptide modulating
 XX flower architecture and flowering time, useful in the area of plant
 XX biotechnology, and commercial plant farming and agriculture.
 XX Disclosure; SEQ ID NO 103; 323pp; English.
 XX The invention comprises the amino acid and coding sequences of proteins
 XX involved in the control of flowering time in rice. The DNA and protein
 XX sequences of the invention are useful for modulating flower architecture
 XX and flowering time, the DNA and protein sequences are useful in the area
 XX of plant biotechnology, commercial plant farming and agriculture. The
 XX present wheat DNA sequence shows homology to the rice sequences of the
 XX invention.
 XX Sequence 1254 BP; 301 A; 296 C; 335 G; 322 T; 0 U; 0 Other;
 XX Query Match 61.8%; Score 20.4; DB 9; Length 1254;
 XX Best Local Similarity 80.0%; Pred. No. 37;
 XX Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGTG 30
 DB 738 CTCTCTCTCTCCGAGCCGGTCGAAATAGTG 767
 RESULT 13
 ID ABT07623/c
 XX ABT07623 standard; cDNA; 698 BP.
 XX AC ABT07623;
 XX 14-NOV-2002 (first entry)
 XX Human breast cancer associated coding sequence SEQ ID NO: 102.
 XX Human; breast specific gene; breast specific protein; breast cancer;
 XX gene therapy; cytostatic; gene; ss.
 XX Homo sapiens.
 XX WO200264611-A1.
 XX 22-AUG-2002.
 XX 12-FEB-2002; 2002WO-US004197.
 XX 13-FEB-2001; 2001US-0268292P.
 XX (DIAD-) DIADEXUS INC.
 XX WPI; 2003-268312/26.
 XX GENBANK; 875037.
 XX New composition comprising two or more isolated polypeptides, useful for
 XX preparing a medicament for treating pain in an animal.
 XX Claim 1; Page: 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 XX or human polynucleotides or a polynucleotide which represents a fragment,
 XX derived or allelic variation of the nucleic acid sequence. Also
 XX claimed are a vector comprising the novel polynucleotide, a host cell
 XX comprising the vector, a method for identifying a nucleotide sequence
 XX which is differentially regulated in an animal subjected to pain and a
 XX kit to perform the method, an array, a method for identifying an agent
 XX that increases or decreases the expression of the polynucleotide sequence
 XX that is differentially expressed in neuronal tissue of a first animal
 XX subjected to pain, a method for identifying a compound which regulates
 XX the expression of a polynucleotide sequence which is differentially
 XX expressed in an animal subjected to pain, a method for identifying a
 XX compound that regulates the activity of one or more of the
 XX polynucleotides, a method for producing a pharmaceutical composition, a
 XX method for identifying a compound or small molecule that regulates the
 XX activity in an animal of one or more of the polypeptides given in the
 XX specification, a method for identifying a compound useful in treating
 XX pain and a pharmaceutical composition comprising the one or more
 XX polypeptides or their antibodies. The polynucleotide or the compound that
 XX modulates its activity is useful for preparing a medicament for treating
 XX pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
 XX therapy). The sequence presented is a human DNA (shown in Table 2 of the
 XX specification) which encodes one of the polypeptides of the invention
 XX which is differentially expressed during pain. Note: The sequence data
 XX for this patent did not form part of the printed specification, but was
 XX obtained in electronic form directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 2695 BP; 772 A; 574 C; 656 G; 693 T; 0 U; 0 Other;
 XX Query Match 61.2%; Score 20.2; DB 9; Length 2695;
 XX Best Local Similarity 75.8%; Pred. No. 52;
 XX Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 CATCTCTCTCCGAGCCGGTCGAAATAGTG 33
 DB 1159 CTCTCTCTCTCCGAGCCGGTCGAAATAGTG 1127
 RESULT 13
 ID ABT07623/c
 XX ABT07623 standard; cDNA; 698 BP.
 XX AC ABT07623;
 XX 14-NOV-2002 (first entry)
 XX Human breast cancer associated coding sequence SEQ ID NO: 102.
 XX Human; breast specific gene; breast specific protein; breast cancer;
 XX gene therapy; cytostatic; gene; ss.
 XX Homo sapiens.
 XX WO200264611-A1.
 XX 22-AUG-2002.
 XX 12-FEB-2002; 2002WO-US004197.
 XX 13-FEB-2001; 2001US-0268292P.
 XX (DIAD-) DIADEXUS INC.

XX PI Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;
 XX PI Sun Y, Liu C;
 XX WPI; 2002-657582/70.
 XX New breast specific nucleic acids and proteins, useful for identifying,
 PT diagnosing, monitoring, staging, imaging, and treating breast cancer and
 PT non-cancerous disease states in breast tissue, and in gene therapy.
 XX
 XX Claim 1; Page 226; 367pp; English.
 XX The present invention provides human breast specific coding sequences and
 CC proteins. These can be used in the diagnosis and treatment of breast
 CC cancer and non-cancerous diseases of the breast. The present sequence is
 CC a coding sequence of the invention
 XX
 XX Sequence 698 BP; 205 A; 126 C; 161 G; 206 T; 0 U; 0 Other;
 SQ
 Query Match 60.6%; Score 20; DB 6; Length 698;
 Best Local Similarity 82.1%; Pred. No. 51;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CATCTCTTCGAGCGCGTGGAAATAG 28
 DB 34 CAGCTCTGCTCCCAAGCGCGTGGAAATGG 7

RESULT 14
 ABK74026/c
 ID ABK74026 standard; DNA; 1018 BP.
 XX AC ABK74026;
 XX
 XX 13-AUG-2002 (first entry)
 XX Bacillus licheniformis genomic sequence tag (GST) #1317.
 XX Differential gene expression; genomic sequenced tag; GST;
 KW altered culture condition; environmental stress;
 KW physiological provocation; ds.
 XX
 XX Bacillus licheniformis.
 XX WO200229113-A2.
 XX
 XX 11-APR-2002.
 XX
 XX 05-OCT-2001; 2001WO-US031437.
 XX 06-OCT-2000; 2000US-00680598.
 XX 27-MAR-2001; 2001US-0279526P.
 XX (NOVO) NOVOZYMES BIOTECH INC.
 XX (NOVO) NOVOZYMES AS.
 XX Berka R, Clausen IG;
 XX WPI; 2002-416684/44.
 XX
 XX Monitoring differential expression of several genes in first Bacillus
 PT cell relative to expression of same genes in one or more second Bacillus
 PT cells, by using substrate containing Bacillus genomic sequenced tag
 PT array.
 XX
 XX Claim 4; SEQ ID NO 1317; 200pp; English.
 XX The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridizing labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of

CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive follow
 CC up characterisation is unnecessary, when one spot on an array equals one
 CC gene or one open reading frame, since sequence information is available.
 CC This sequence represents a genomic sequence tag (GST) used in the method
 CC of the invention. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences;
 XX

SQ Sequence 1018 BP; 328 A; 207 C; 266 G; 217 T; 0 U; 0 Other;
 Query Match 60.6%; Score 20; DB 6; Length 1018;
 Best Local Similarity 82.1%; Pred. No. 54;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 2 ATCTCTCTCCGAGCGCGTGGAAATAGT 29
 DB 845 ATGACTCTCCCGTGGCGTGGAAATAGT 818

RESULT 15
 ABT07624/c
 ID ABT07624 standard; cDNA; 1217 BP.
 XX AC ABT07624;
 XX
 XX 14-NOV-2002 (first entry)
 XX Human breast cancer associated coding sequence SEQ ID NO: 103.
 XX Human breast specific gene; breast specific protein; breast cancer;
 KW gene therapy; cytostatic; gene; ss.
 XX Homo sapiens.
 XX WO200264611-A1.
 XX
 XX 22-AUG-2002.
 XX
 XX 12-FEB-2002; 2002WO-US004197.
 XX 13-FEB-2001; 2001US-0268292P.
 XX (DIAD-) DIADEXUS INC.
 XX Salceda S, Macina RA, Hu P, Recipon H, Karra K, Cafferkey R;
 XX Sun Y, Liu C;
 XX WPI; 2002-657582/70.
 XX New breast specific nucleic acids and proteins, useful for identifying,
 PT diagnosing, monitoring, staging, imaging, and treating breast cancer and
 PT non-cancerous disease states in breast tissue, and in gene therapy.
 XX
 XX Claim 1; Page 226-227; 367pp; English.
 XX The present invention provides human breast specific coding sequences and
 CC proteins. These can be used in the diagnosis and treatment of breast
 CC cancer and non-cancerous diseases of the breast. The present sequence is
 CC a coding sequence of the invention
 XX
 XX Sequence 1217 BP; 385 A; 210 C; 277 G; 345 T; 0 U; 0 Other;
 SQ
 Query Match 60.6%; Score 20; DB 6; Length 1217;
 Best Local Similarity 82.1%; Pred. No. 56;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CATCTCTCTCCGAGCCGGTCGAAATAG 28
 |||||
 Db 34 CAGCTCTCTCCAGCCCGTCGAAATGG 7

RESULT 16
 AAV30458_4
 Continuation (5 of 6) of AAV30458 from base 400001 (Rhisobium species plasmid pNGR234a.
 WP Sequence split into 6 fragments LOCUS AAV30458 Accession AAV30458

Fragment Name	Begin	End
AAV30458_0	1	110000
AAV30458_1	100001	210000
AAV30458_2	200001	310000
AAV30458_3	300001	410000
AAV30458_4	400001	510000
AAV30458_5	500001	534720

Query Match 60.6%; Score 20; DB 2; Length 110000;
 Best Local Similarity 52.1%; Pred. No. 1.1e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CTCCTCTCCGAGCCGGTCGAAATAGTGA 31
 |||||
 Db 55880 CTCCTCAGGAGCCGGTCGACCTGCTGA 55907

RESULT 17
 AAV30459_4
 Continuation (5 of 6) of AAV30459 from base 400001 (Rhisobium species symbiotic plasmid
 WP Sequence split into 6 fragments LOCUS AAV30459 Accession AAV30459

Fragment Name	Begin	End
AAV30459_0	1	110000
AAV30459_1	100001	210000
AAV30459_2	200001	310000
AAV30459_3	300001	410000
AAV30459_4	400001	510000
AAV30459_5	500001	536165

Query Match 60.6%; Score 20; DB 2; Length 110000;
 Best Local Similarity 52.1%; Pred. No. 1.1e+02;
 Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CTCCTCTCCGAGCCGGTCGAAATAGTGA 31
 |||||
 Db 55880 CTCCTCAGGAGCCGGTCGACCTGCTGA 55907

RESULT 18
 AAC40134
 ID AAC40134 standard; DNA; 1230 BP.

XX AAC40134;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SRQ ID NO: 27179.

XX Hybridization assay; genetic mapping; gene expression control;
 XX protein identification; signal transduction pathway; metabolic pathway;
 XX promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EF1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 200CEP-00301439.

XX 25-FEB-1999; 99US-0121805P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999;	99US-0125788P.
PR 25-MAR-1999;	99US-0126264P.
PR 29-MAR-1999;	99US-0126785P.
PR 01-APR-1999;	99US-0127462P.
PR 06-APR-1999;	99US-0128231P.
PR 08-APR-1999;	99US-0128714P.
PR 16-APR-1999;	99US-0129645P.
PR 19-APR-1999;	99US-0130077P.
PR 21-APR-1999;	99US-0130449P.
PR 23-APR-1999;	99US-0130510P.
PR 23-APR-1999;	99US-0130891P.
PR 28-APR-1999;	99US-0131449P.
PR 30-APR-1999;	99US-0132048P.
PR 30-APR-1999;	99US-0132407P.
PR 04-MAY-1999;	99US-0132484P.
PR 05-MAY-1999;	99US-0132485P.
PR 06-MAY-1999;	99US-0132485P.
PR 06-MAY-1999;	99US-0132487P.
PR 07-MAY-1999;	99US-0132863P.
PR 11-MAY-1999;	99US-013256P.
PR 14-MAY-1999;	99US-013218P.
PR 14-MAY-1999;	99US-013219P.
PR 14-MAY-1999;	99US-013221P.
PR 14-MAY-1999;	99US-0132370P.
PR 18-MAY-1999;	99US-0132768P.
PR 19-MAY-1999;	99US-0132941P.
PR 20-MAY-1999;	99US-0132124P.
PR 21-MAY-1999;	99US-0132353P.
PR 24-MAY-1999;	99US-0132623P.
PR 25-MAY-1999;	99US-0136021P.
PR 27-MAY-1999;	99US-0136392P.
PR 28-MAY-1999;	99US-0136782P.
PR 01-JUN-1999;	99US-0137222P.
PR 03-JUN-1999;	99US-0137528P.
PR 04-JUN-1999;	99US-0137502P.
PR 07-JUN-1999;	99US-0137724P.
PR 08-JUN-1999;	99US-0138094P.
PR 10-JUN-1999;	99US-0138540P.
PR 10-JUN-1999;	99US-0138647P.
PR 14-JUN-1999;	99US-0139113P.
PR 16-JUN-1999;	99US-0139452P.
PR 16-JUN-1999;	99US-0139453P.
PR 17-JUN-1999;	99US-0139492P.
PR 18-JUN-1999;	99US-0139454P.
PR 18-JUN-1999;	99US-0139455P.
PR 18-JUN-1999;	99US-0139456P.
PR 18-JUN-1999;	99US-0139457P.
PR 18-JUN-1999;	99US-0139458P.
PR 18-JUN-1999;	99US-0139459P.
PR 18-JUN-1999;	99US-0139460P.
PR 18-JUN-1999;	99US-0139461P.
PR 18-JUN-1999;	99US-0139462P.
PR 18-JUN-1999;	99US-0139463P.
PR 18-JUN-1999;	99US-0139750P.
PR 18-JUN-1999;	99US-0139763P.
PR 21-JUN-1999;	99US-0139817P.
PR 22-JUN-1999;	99US-0139899P.
PR 23-JUN-1999;	99US-0140353P.
PR 23-JUN-1999;	99US-0140354P.
PR 24-JUN-1999;	99US-0140695P.
PR 28-JUN-1999;	99US-0140823P.
PR 29-JUN-1999;	99US-0140991P.
PR 30-JUN-1999;	99US-0141287P.
PR 01-JUL-1999;	99US-0141842P.
PR 02-JUL-1999;	99US-0142154P.
PR 06-JUL-1999;	99US-0142055P.
PR 08-JUL-1999;	99US-0142390P.
PR 09-JUL-1999;	99US-0142803P.
PR 12-JUL-1999;	99US-0142920P.
PR 13-JUL-1999;	99US-0142977P.
PR 13-JUL-1999;	99US-0143542P.
PR 14-JUL-1999;	99US-0143624P.
PR 15-JUL-1999;	99US-0144005P.

XX Hearing P, Bahou WP, Sandalon Z, Gnatenko DV;
 XX WPI, 2002-690619/74.
 XX
 XX Producing vector, by introducing vector having nucleotide sequence,
 PT adenovirus inverted terminal repeats and packaging sequence, and adeno-
 PT associated virus terminal repeat, into cell, and culturing cell.
 XX
 XX Disclosure; Page 166-181; 131pp; English.
 XX
 XX The present invention relates to a new method of producing a vector. The
 CC method involves introducing recombinant vector having nucleotide sequence
 CC (NS) having 5' and 3' end, left and right inverted terminal repeats of
 CC adenovirus flanking NS, adenovirus packaging sequence linked to inverted
 CC terminal repeat, and adeno-associated virus terminal repeat linked to 3'
 CC end of NS, into cell expressing adenovirus early gene lacking from vector
 CC ; and culturing cell to produce another vector. The method is useful for
 CC generating vectors, especially mad vectors. The method is useful in
 CC transferring nucleotide sequences of interest into a cell, for gene
 CC transfer applications (e.g. gene therapy) in vitro, ex vivo and in vivo.
 CC The nucleotide sequences are useful for treating diseases associated with
 CC it, i.e. adenosine deaminase gene associated with adenosine deaminase
 CC deficiency with severe combined immune deficiency, beta-chain of
 CC haemoglobin gene associated with beta-thalassemia and sickle cell
 CC disease, receptor for low density lipoprotein gene associated with
 CC familial hypercholesterolemia, hypoxanthine-guanine
 CC phosphoribosyltransferase associated with Lesch-Nyhan syndrome,
 CC phenylalanine hydroxylase (PAH) gene associated with phenylketonuria,
 CC dytrophin gene associated with muscular dystrophy, and human cystic
 CC fibrosis transmembrane conductance regulator gene associated with cystic
 CC fibrosis. The present nucleic acid sequence represents a human adenovirus
 CC genome sequence that was used in the methods of the invention
 XX
 XX Sequence 34214 BP; 8509 A; 8969 C; 8556 G; 8180 T; 0 U; 0 Other;
 SQ
 Query Match 58.8%; Score 19.4; DB 6; Length 34214;
 Best Local Similarity 79.3%; Pred. No. 1.8e+03;
 Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 2 ATCTCTTCTCCGAGCGGTGCGAATAGTG 30
 DB 12904 ATTTTTCCTCCGCGCGGTGCGAATAGG 12876
 RESULT 20
 AAL45392
 ID AAL45392 standard; DNA; 50 BP.
 XX
 XX AAL45392,
 AC
 XX
 XX 06-JUN-2002 (first entry)
 DT
 XX
 XX Zinc-dependent deoxyribozyme SEQ ID NO: 14.
 XX
 XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
 KW quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.
 KW
 XX Unidentified.
 OS
 XX WO200200006-A2.
 XX
 XX 03-JAN-2002.
 PD
 XX
 XX 27-JUN-2001; 2001WO-US020557.
 PP
 XX 27-JUN-2000; 2000US-0060558.
 XX
 XX (UNII) UNIV ILLINOIS FOUND.
 PA
 XX Lu Y, Li J;
 F-
 XX WPI; 2002-130823/17.
 DR

XX New nucleic acid enzyme biosensors, useful for the sensitive and
 PT selective detection of ions, particularly metal ions e.g. lead ions, and
 PT for determining the concentration of a particular ion in a solution.
 XX
 XX Example 1; Fig 2; 57pp; English.
 PS
 XX The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependent on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a ribozyme
 CC which may be used as a biosensor of the invention
 XX
 XX Sequence 50 BP; 12 A; 10 C; 12 G; 16 T; 0 U; 0 Other;
 SQ
 Query Match 58.2%; Score 19.2; DB 6; Length 50;
 Best Local Similarity 75.0%; Pred. No. 78;
 Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 2 ATCTCTTCTCCGAGCGGTGCGAATAGTGAGT 33
 DB 1 ATCTCTTCTCCGAGCGGTGCGAATAGTGAGT 32
 RESULT 21
 ABL30213/c
 ID ABL30213 standard; DNA; 1512 BP.
 XX
 XX ABL30213;
 AC
 XX 26-MAR-2002 (first entry)
 DT
 XX
 XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 42112.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ds.
 KW
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 PD
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 PP
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 XX 11-JUL-2000; 2000US-00614150.
 XX
 XX (PFXE) PG CORP NY.
 PA
 XX
 XX Venter JC, Adams M, Li PMD, Myers EW;
 PI
 XX WPI, 2001-656860/75.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 PT
 XX Claim 1; SEQ ID NO 42112; 21pp + Sequence Listing; English.
 PS
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL18175) and the encoded proteins (ABR57737
 CC ABR7072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

```

XX SQ Sequence 1512 BP; 367 A; 403 C; 402 G; 340 T; 0 U; 0 Other;
XX Query Match 57.6%; Score 19; DB 4; Length 1512;
XX Best Local Similarity 81.5%; Pred. No. 1.9e+02;
XX Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CATCTCTCTCCGAGCGCGTGGAAATA 27
DB 85 CATTTCTACTCCGAGCGCGTTAAATA 59

RESULT 22
ABL30212
ID ABL30212 standard; DNA; 3686 BP.
XX AC ABL30212;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 42109.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX PD 27-SEP-2001.
XX PF 23-MAR-2001; 2001WO-US009231.
XX PR 23-MAR-2000; 2000US-0191637P.
XX PR 11-JUL-2000; 2000US-00614150.
XX PA (PEKE) PE CORP NY.
XX PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
XX Claim 1, SEQ ID NO 42109; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL03840-ABL16175) and the encoded proteins (ABBS7737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 3686 BP; 993 A; 865 C; 837 G; 991 T; 0 U; 0 Other;
XX Query Match 57.6%; Score 19; DB 4; Length 3686;
XX Best Local Similarity 81.5%; Pred. No. 1.9e+02;
XX Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 CATCTCTCTCCGAGCGCGTGGAAATA 27
DB 2602 CATTTCTACTCCGAGCGCGTTAAATA 2628

RESULT 23
AXX88481/c
ID AXX88481 standard; DNA; 50 BP.

```

```

XX AXX88481;
XX AC 01-OCT-1999 (first entry)
XX DT Human MIP-1 beta primer PIC*.
XX DE RANTES; chemokine; detection; primer; probe; amplification; MIP-1 alpha;
XX KW regulated upon activation normal T expressed and secreted; MIP-1 beta;
XX KW macrophage inflammatory protein; CD4+ T-cell; inhibitor; prognosis;
XX KW primary non-synctium-inducing HIV-1 strain; therapy; ss.
XX OS Synthetic.
XX OS Homo sapiens.
XX PN WO9917815-A1.
XX PD 29-JUL-1999.
XX PF 22-JAN-1999; 99WO-US001327.
XX PR 22-JAN-1998; 98US-00010641.
XX PA (ALKU) AKZO NOBEL NV.
XX PI Romano JW, Shurtliff R, Williams KG;
XX WPI; 1999-459145/39.
XX Detection of expression levels of the cytokines RANTES, MIP-1alpha and
XX MIP-1beta used as prognostic markers of HIV-infected patients.
XX Claim 17b; Page 43; 48pp; English.
XX This invention describes novel oligonucleotides which are used for
XX detecting the chemokines RANTES (regulated upon activation normal T
XX expressed and secreted), macrophage inflammatory protein (MIP)-1 alpha or
XX MIP-1 beta by (a) obtaining a sample possible containing RANTES or MIP-1
XX alpha or MIP-1 beta RNA, (b) performing an isothermal transcriptional
XX amplification on the sample with 2 oligonucleotide primers, (c) detecting
XX the product of step (b) where detection of a product indicates the
XX presence of RANTES, MIP-1 alpha or MIP-1 beta in the sample. The assay is
XX used to determine the levels of the chemokines RANTES, MIP-1 alpha and
XX MIP-1 beta in samples, especially cells. These chemokines have been shown
XX to be inhibitors of CD4+ T-cells by primary non-synctium-inducing HIV-1
XX strains. Thus the level of expression of these genes can be used as
XX prognostic markers for direct therapeutic management of HIV-infected
XX patients. By being isothermal, the assay requires less manipulation by
XX the experimenter. Also, spiking the sample with a known amount of
XX control RNA allows quantitation and qualification of the products in a
XX single assay. AXX88481-XX88491 represent the primers and probes used in the
XX method of the invention
XX SQ Sequence 50 BP; 12 A; 10 C; 17 G; 11 T; 0 U; 0 Other;
XX Query Match 57.0%; Score 18.8; DB 2; Length 50;
XX Best Local Similarity 76.7%; Pred. No. 1.2e+02;
XX Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CTCTCTCTCCGAGCGCGTGGAAATAGTGA 33
DB 42 CTCTCTCTCCGAGCGCGTGGAAATAGTGA 13

RESULT 24
AAL45395
ID AAL45395 standard; DNA; 51 BP.
XX AC AAL45395;
XX DT 06-JUN-2002 (first entry)
XX DE Zinc-dependent deoxyribozyme SEQ ID NO: 17.

```


XX New isolated *Trypanosoma pallidum* nucleic acids - used to develop products
 PT for the detection, diagnosis, characterization, prevention and therapy of
 PT *T. pallidum* infections, particularly syphilis.

XX Claim 1, Page 900; 1150pp; English.

XX AAX20500-21243 represent polynucleotide sequences from the genome of
 CC *Trypanosoma pallidum*. The sequences can be used for detection, diagnosis,
 CC characterization, prevention and therapy for *T. pallidum* infections,
 CC particularly syphilis. They can also be used for detecting diseases
 CC related to *Borrelia* infections in animals, and for the production of
 CC biosynthetic products such as enzymes

XX SQ Sequence 636 BP; 175 A; 163 C; 136 G; 161 T; 0 U; 1 Other;

Query Match 57.0%; Score 18.8; DB 2; Length 636;
 Best Local Similarity 76.7%; Pred. No. 1.8e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 4 CTCTCTCCGAGCGGTCGAAATAGTGGT 33

DB 182 CACATCTCCGAGCTGGTCCAGTAGTGGGT 153

RESULT 27

ID ADA49038 standard; DNA; 734 BP.

XX AC ADA49038;

XX DT 20-NOV-2003 (first entry)

XX DB Wheat gene conferring disease resistance in plants.

XX disease resistance; pathogen tolerance; plant pathogen; ds; gene; plant;
 XX wheat.

XX OS Triticum aestivum.

XX FN WO2003000906-A2.

XX PD 03-JAN-2003.

XX PF 21-JUN-2002; 2002WO-18002453.

XX PR 22-JUN-2001; 2001US-0300112P.

XX PR 26-SEP-2001; 2001US-0352277P.

XX PR 22-MAR-2002; 2002US-0366535P.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T,
 XX PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;

XX DR WPI; 2003-184052/18.

XX New polynucleotide comprising a plant nucleotide sequence having an open
 PT reading frame that encodes a polypeptide associated with disease
 PT resistance, useful for conferring resistance or tolerance to a plant
 PT pathogen.

XX DS Disclosure; SEQ ID NO 1108; 299pp; English.

XX The invention relates to a novel isolated polynucleotide comprising a
 CC plant nucleotide sequence having an open reading frame that encodes a
 CC polypeptide associated with disease resistance or its fragment having
 CC substantially the same activity as the full-length polypeptide. The
 CC polynucleotide of the invention is useful for conferring resistance or
 CC tolerance to a plant pathogen. The present sequence represents a gene
 CC conferring disease resistance used in the invention.

XX SQ Sequence 734 BP; 166 A; 207 C; 222 G; 139 T; 0 U; 0 Other;

Query Match 57.0%; Score 18.8; DB 8; Length 734;
 Best Local Similarity 76.7%; Pred. No. 1.8e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 3 TCTCTCTCCGAGCGGTCGAAATAGTGGT 32

DB 555 TCTCTCCAGCGGCGCTAGAAATGGG 526

RESULT 28

ID ADC03443/c

XX AC ADC03443 standard; DNA; 734 BP.

XX AC ADC03443;

XX DT 18-DEC-2003 (first entry)

XX DB Wheat flowering time-related DNA sequence #3.

XX flowering time; wheat; flower architecture; plant biotechnology;
 XX commercial plant farming; agriculture; flowering-related protein; ds.

XX OS Triticum aestivum.

XX FN WO2003000904-A2.

XX PD 03-JAN-2003.

XX PF 24-JUN-2002; 2002WO-EP006968.

XX PR 22-JUN-2001; 2001US-0300112P.

XX PR 26-SEP-2001; 2001US-0352277P.

XX PR 30-NOV-2001; 2001US-0334984P.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Sessions A, Briggs S, Cooper B, Goff SA, Moughamer T,
 XX PI Glazebrook J, Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;

XX DR WPI; 2003-229340/22.

XX New isolated nucleic acid molecule encoding a polypeptide modulating
 PT flower architecture and flowering time, useful in the area of plant
 PT biotechnology, and commercial plant farming and agriculture.

XX PS Disclosure; SEQ ID NO 83; 323pp; English.

XX The invention comprises the amino acid and coding sequences of proteins
 CC involved in the control of flowering time in rice. The DNA and protein
 CC sequences of the invention are useful for modulating flower architecture
 CC and flowering time, the DNA and protein sequences are useful in the area
 CC of plant biotechnology, commercial plant farming and agriculture. The
 CC present wheat DNA sequence shows homology to the rice sequences of the
 CC invention.

XX SQ Sequence 734 BP; 166 A; 207 C; 222 G; 139 T; 0 U; 0 Other;

Query Match 57.0%; Score 18.8; DB 9; Length 734;
 Best Local Similarity 76.7%; Pred. No. 1.8e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 3 TCTCTCTCCGAGCGGTCGAAATAGTGGT 32

DB 555 TCTCTCCAGCGGCGCTAGAAATGGG 526

RESULT 29

ID ACA26446

XX AC ACA26446 standard; DNA; 846 BP.

XX AC ACA26446;

XX

DT XX 19-JUN-2003 (first entry)
 DE XX Prokaryotic essential gene #8103.
 XX Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX Burkholderia mallei.
 OS WO200277183-A2.
 PN 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948991.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU22576.
 DR New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids, required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 14; SEQ ID NO 14316; 1766pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *X. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target
 CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 846 BP; 153 A; 286 C; 127 T; 0 U; 0 Other;
 Query Match 57.0%; Score 18.8; DB 7; Length 846;
 Best Local Similarity 76.7%; Pred. No. 1.8e+03;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ATCTCTTCTCCGAGCCGGTCCGAATAGTGA 31
 DB 761 ATCCGTACGACAAAGCCGCTCCGATCGTGA 790
 RESULT 30
 ACA26384
 ID ACA26384 standard; DNA; 1371 BP.
 XX AC ACA26384;
 XX 19-JUN-2003 (first entry)
 XX Prokaryotic essential gene #8041.
 KW Antisense; ds; prokaryotic essential gene; cell proliferation;
 KW drug design; gene.
 XX Burkholderia mallei.
 OS WO200277183-A2.
 PN 03-OCT-2002.
 XX 21-MAR-2002; 2002WO-US009107.
 XX 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948991.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR P-PSDB; ABU22514.
 DR New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids, required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX Claim 14; SEQ ID NO 14254; 1766pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *X. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target

CC prokaryotic essential genes. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 1371 BP; 138 A; 496 C; 484 G; 253 T; 0 U; 0 Other;
 Query Match 57.0%; Score 18.8; DB 7; Length 1371;
 Best Local Similarity 90.9%; Pred. No. 2e+02;
 Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 ATCTCTTCTCCGAGCGCGTGA 23
 DB 380 ATCTCTTCTCCGAGCGCGGCA 401

RESULT 31
 ABQ61114
 ID ABQ61114 standard; cDNA; 1926 BP.
 XX AC ABQ61114;
 XX DT 26-FEB-2003 (first entry)
 XX DE FLJ20123 f1s clone encoding sequence.
 XX KM Neuroprotective; immunomodulator; cancer; chromosome 17; cytostatic;
 XX KN anti-inflammatory; gene therapy; nutritional supplement; wound; burn;
 XX KW ulcer; Alzheimer's disease; Huntington's disease;
 XX KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 XX KW vulnery; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200231111-A2.
 XX PD 18-APR-2002.
 XX PF 11-OCT-2001; 2001WO-US027760.
 XX PR 12-OCT-2000; 2000US-00687527.
 XX PA (HYSE-) HYSHQ INC.
 XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 XX PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX DR WPI, 2002-426278/45.
 XX DR N-PSDB; ABP43870.
 XX PT New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.
 XX PS Claim 1; SEQ ID # 327; 357pp + Sequence Listing; English.

XX The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC vulnery, neuroprotective, immunomodulator, cytostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records ABQ60788-
 CC ABQ61233 represent polynucleotides of the invention. NOTE: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pat_sequences

XX SQ Sequence 1926 BP; 384 A; 583 C; 521 G; 438 T; 0 U; 0 Other;

Query Match 57.0%; Score 18.8; DB 6; Length 1926;
 Best Local Similarity 76.7%; Pred. No. 2.1e+02;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 CACTCTCTCTCCGAGCGCGTGAATAGT 30
 DB 1711 CATTGCTTCTCTGAGCGCTCTCTGACATG 1740

RESULT 32
 AAL45394
 ID AAL45394 standard; DNA; 49 BP.
 XX AC AAL45394;
 XX DT 06-JUN-2002 (first entry)
 XX DE Zinc-dependent deoxyribozyme SEQ ID NO: 16.
 XX KM Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
 XX KW quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.
 XX OS Unidentified.
 XX PN WO200200006-A2.
 XX PD 03-JAN-2002.
 XX PF 27-JUN-2001; 2001WO-US020557.
 XX PR 27-JUN-2000; 2000US-00605558.
 XX PA (UNII) UNIV ILLINOIS FOUND.
 XX PI Lu Y, Li J;
 XX DR WPI, 2002-130823/17.
 XX PT New nucleic acid enzyme biosensors, useful for the sensitive and
 PT selective detection of ions, particularly metal ions e.g. lead ions, and
 PT for determining the concentration of a particular ion in a solution.
 XX PS Example 1, Fig 2; 57pp; English.

XX The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependent on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a ribozyme
 CC which may be used as a biosensor of the invention

XX SQ Sequence 49 BP; 13 A; 10 C; 12 G; 14 T; 0 U; 0 Other;
 Query Match 56.4%; Score 18.6; DB 6; Length 49;
 Best Local Similarity 84.0%; Pred. No. 1.4e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 7 TTCTCCGAGCGCGTCCGAATAGTGA 31
 DB 9 TTCTCCGAGCGCGTCCGAATAGTGA 33

RESULT 33
 AAL45393
 ID AAL45393 standard; DNA; 49 BP.
 XX AC AAL45393;
 XX DT 06-JUN-2002 (first entry)
 XX

PA (HYBR-) HYBRIGENICS.
 PA (INSP) INST PASTEUR.
 PI Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
 XX WPI; 2002-674910/72.
 DR P-PSDB; ABUS0898.
 DR
 XX New complexes of protein-protein interactions in Helicobacter pylori,
 PT useful for identifying modulating compounds for treating or preventing
 PT ulcers in mammals.
 XX
 XX Claim 7; Page 144; 642pp; English.
 XX
 CC The invention describes a complex of protein-protein interactions in
 CC Helicobacter pylori selected from 421 complexes given in the
 CC specification. The complex of protein-protein interactions are useful for
 CC screening for agents which modulate the interaction of proteins.
 CC Modulating compounds which binds to a targeted bacterial protein may be
 CC used for treating or preventing ulcers in a human or animal. This
 CC sequence encodes a selected interacting domain (SID), identified via
 CC protein-protein interactions
 CC
 CC Sequence 1024 BP; 267 A; 186 C; 240 G; 331 T; 0 U; 0 Other;
 XX
 Query Match 56.4%; Score 18.6; DB 6; Length 1024;
 Best Local Similarity 84.0%; Pred. No. 2.3e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 XX
 QY 1 CATCTCTCTCCGAGCGGTCGAAA 25
 Db 855 CATCTCTCTCTGATCCTGTCGAAA 831
 XX
 RESULT 39
 AEX66280/C
 ID AEX66280 standard; DNA; 1024 BP.
 AC AEX66280;
 XX
 DT 07-MAY-2003 (first entry)
 DE Helicobacter pylori selected interacting domain (SID) DNA #879.
 XX
 XX Protein-protein interaction; ulcer; selected interacting domain; SID;
 XX gene; ds.
 XX
 OS Helicobacter pylori.
 XX
 XX WO200266501-A2.
 XX
 XX 29-AUG-2002.
 XX
 XX 28-DEC-2001; 2001WO-EP015428.
 XX
 XX 02-JAN-2001; 2001US-0259302P.
 XX
 PA (HYBR-) HYBRIGENICS.
 PA (INSP) INST PASTEUR.
 XX
 XX Legrain P, Rain J, Colland F, De Reuse H, Labigne A;
 XX WPI; 2002-674910/72.
 DR P-PSDB; ABUS1536.
 DR
 XX New complexes of protein-protein interactions in Helicobacter pylori,
 PT useful for identifying modulating compounds for treating or preventing
 PT ulcers in mammals.
 XX
 XX Claim 7; Page 298; 642pp; English.
 XX
 CC The invention describes a complex of protein-protein interactions in
 CC Helicobacter pylori selected from 421 complexes given in the

CC specification. The complex of protein-protein interactions are useful for
 CC screening for agents which modulate the interaction of proteins.
 CC Modulating compounds which binds to a targeted bacterial protein may be
 CC used for treating or preventing ulcers in a human or animal. This
 CC sequence encodes a selected interacting domain (SID), identified via
 CC protein-protein interactions
 CC
 CC Sequence 1024 BP; 267 A; 186 C; 240 G; 331 T; 0 U; 0 Other;
 XX
 Query Match 56.4%; Score 18.6; DB 6; Length 1024;
 Best Local Similarity 84.0%; Pred. No. 2.3e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 XX
 QY 1 CATCTCTCTCCGAGCGGTCGAAA 25
 Db 855 CATCTCTCTCTGATCCTGTCGAAA 831
 XX
 RESULT 40
 AAX14487/C
 ID AAX14487 standard; DNA; 1035 BP.
 XX
 AC AAX14487;
 XX
 DT 31-MAR-1999 (first entry)
 DE H. pylori GHPO 1125 gene.
 XX
 XX GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
 XX peptic ulcer disease; ss.
 XX
 OS Helicobacter pylori.
 XX
 XX Key Location/Qualifiers
 XX CDS 25..996
 XX FT /tag= a
 XX
 XX WO9843478-A1.
 XX
 XX 08-OCT-1998.
 XX
 XX 01-APR-1998; 98WO-USC06371.
 XX
 XX 01-APR-1997; 97US-00833457.
 XX
 XX 24-JUN-1997; 97US-00861227.
 XX
 XX 29-JUL-1997; 97US-00902615.
 XX
 XX (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
 XX (HUNA-) HUMAN GENOME SCI INC.
 XX
 XX Kleanthous H, Al-Garawi A, Miller C, Tomb J, Oomen RP;
 XX WPI; 1998-542293/46.
 XX P-PSDB; AAW98768.
 XX
 XX New isolated Helicobacter polynucleotides - used to develop products for
 XX the diagnosis, prevention and treatment of Helicobacter infections and
 XX gastrointestinal diseases.
 XX
 XX Claim 1; Page 1658-1659; 2054pp; English.
 XX
 CC This sequence represents a polynucleotide of the invention. It was
 CC isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.
 CC The polypeptides can be used for preventing or treating Helicobacter
 CC infections, and gastroduodenal diseases associated with these infections,
 CC including acute, chronic, and atrophic gastritis, and peptic ulcer
 CC diseases, e.g. gastric and duodenal ulcers. They can also be used for the
 CC production of antibodies. The products can also be used for detection and
 CC diagnosis
 XX
 XX Sequence 1035 BP; 281 A; 168 C; 256 G; 329 T; 0 U; 1 Other;
 XX
 Query Match 56.4%; Score 18.6; DB 2; Length 1035;

Best Local Similarity 84.0%, Pred. No. 2.3e+02;
Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 CATCTTTCTCCGAGCCGGTCGAAA 25
DB 672 CATCTTTCTCTGATCCTGTCAAAA 648

Search completed: May 24, 2004, 11:14:44
Job time : 206.981 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 10:18:57 ; Search time 973.189 Seconds
(without alignments)
1469.725 Million cell updates/sec

Title: US-10-144-679-1

Perfect score: 33

Sequence: 1 catctctctccgagccgctgcaaatgtagt 33

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 5940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pri.*
10: gb_ro.*
11: gb_sce.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mus.*
20: em_on.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sta.*
28: em_un.*
29: em_vi.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pln.*
35: em_hg_rdd.*
36: em_hg_mam.*
37: em_hg_vrt.*
38: em_ey.*
39: em_hggo_hum.*
40: em_hggo_mus.*
41: em_hggo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	33	100.0	33	6	AX18515	AX18515 Sequence
2	33	100.0	33	6	AX18536	AX18536 Sequence
3	27.8	84.2	50	6	AX18534	AX18534 Sequence
4	26.4	80.0	50	6	AX18532	AX18532 Sequence
5	26.4	80.0	50	6	AX18533	AX18533 Sequence
6	24	72.7	50	6	AX18535	AX18535 Sequence
7	21.8	66.1	33	6	AX18521	AX18521 Sequence
8	21.4	64.8	237565	2	AC112761	AC112761 Rattus no
9	21.4	64.8	234801	2	AC118109	AC118109 Rattus no
10	21.4	64.8	244180	2	AC094031	AC094031 Rattus no
11	21	63.6	3356	8	SCU20322	U20322 Saccharomyc
12	20.6	62.4	475	6	AX194850	AX194850 Sequence
13	20.6	62.4	1248	6	AX467748	AX467748 Sequence
14	20.6	62.4	1338	10	MUSGAC	M74517 Mouse GA bl
15	20.6	62.4	2598	10	MUSGAB	M74516 Mouse GA bl
16	20.6	62.4	2622	10	BC013558	BC013558 Mus muscu
17	20.6	62.4	122211	10	AL844555	AL844555 Mouse DNA
18	20.6	62.4	141132	2	AC008684	AC008684 Homo sapi
19	20.6	62.4	195909	9	AF165124	AF165124 Homo sapi
20	20.6	62.4	209816	9	AC091926	AC091926 Homo sapi
21	20.4	61.8	1254	6	AX558810	AX558810 Sequence
22	20.4	61.8	17683	10	AF320617	AF320617 Mus muscu
23	20.4	61.8	248911	2	AC109717	AC109717 Rattus no
24	20.4	61.8	312050	1	MLBPRN3	AL583919 Mycobacte
25	20.4	61.8	345783	1	AP003001	AP003001 Mesorhizo
26	20.2	61.2	61580	2	AC100516	AC100516 Mus muscu
27	20.2	61.2	161955	9	AC068718	AC068718 Homo sapi
28	20.2	61.2	177264	10	AL626782	AL626782 Mouse DNA
29	20.2	61.2	207051	2	AC110030	AC110030 Mus muscu
30	20.2	61.2	231904	10	AC122907	AC122907 Mus muscu
31	20.2	61.2	348050	1	AP003581	AP003581 Noctoc ap
32	20	60.6	1018	6	AX432802	AX432802 Sequence
33	20	60.6	1122	6	AR185706	AR185706 Sequence
34	20	60.6	10140	1	AB000105	AB000105 Rhizobium
35	20	60.6	10242	1	AB000102	AB000102 Rhizobium
36	20	60.6	42430	3	CBRG44D18	AC084512 Caenorhab
37	20	60.6	53371	2	AC084186	AC084186 Homo sapi
38	20	60.6	84743	2	AC140870	AC140870 Homo sapi
39	20	60.6	85952	9	HS4373N24	AL121932 Human DNA
40	20	60.6	118767	9	HS4PTBL	Z95704 Human DNA s
41	20	60.6	122652	2	AC138997	AC138997 Homo sapi
42	20	60.6	171310	2	AL354817	AL354817 Homo sapi
43	20	60.6	176995	9	AL471879	AL471879 Human DNA
44	20	60.6	226349	2	AC140865	AC140865 Homo sapi
45	20	60.6	236165	6	A79351	A79351 Sequence 2

ALIGNMENTS

RESULT 1	AX18515	Sequence 1 from Patent WO0100006.	33 bp	DNA	linear	PAT 18-JUN-2002
AX18515	Sequence 1 from Patent WO0100006.					
LOCUS	AX18515					
DEFINITION	AX18515					
ACCESSION	AX18515.1	GI:21523380				
VERSION	AX18515.1	GI:21523380				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						

THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

FEATURES
source
Location/Qualifiers
1. .33
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Trans-cleaving deoxyribozyme 178"

ORIGIN

Query Match 100.0%; Score 33; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATCTCTTCCGAGCCGCTCGAATAAGTAGT 33
|||||
Db 1 CATCTCTTCCGAGCCGCTCGAATAAGTAGT 33
|||||

RESULT 2
AX418596
LOCUS AX418596 33 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 82 from Patent WO200006.
ACCESSION AX418596
VERSION AX418596.1 GI:21523461
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Lu, Y. and Li, J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 020006-A 82 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

FEATURES
source

Location/Qualifiers
1. .33
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Combined DNA/RNA Molecule: Chimeric substrate"

ORIGIN

Query Match 100.0%; Score 33; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CATCTCTTCCGAGCCGCTCGAATAAGTAGT 33
|||||
Db 1 CATCTCTTCCGAGCCGCTCGAATAAGTAGT 33
|||||

RESULT 3
AX418534
LOCUS AX418534 50 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 20 from Patent WO200006.
ACCESSION AX418534
VERSION AX418534.1 GI:21523399
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Lu, Y. and Li, J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 020006-A 20 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

FEATURES
source

Location/Qualifiers
1. .50
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Zn-DNA"

ORIGIN

Query Match 84.2%; Score 27.8; DB 6; Length 50;
Best Local Similarity 93.8%; Pred. No. 0.28;
Matches 29; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ATCTCTTCCGAGCCGCTCGAATAAGTAGT 32
|||||
Db 1 ATCTCTTCCGAGCCGCTCGAATAAGTAGT 31
|||||

RESULT 4
AX418532
LOCUS AX418532 50 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 18 from Patent WO200006.
ACCESSION AX418532
VERSION AX418532.1 GI:21523397
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Lu, Y. and Li, J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 020006-A 18 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

FEATURES
source

Location/Qualifiers
1. .50
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Zn-DNA"

ORIGIN

Query Match 80.0%; Score 26.4; DB 6; Length 50;
Best Local Similarity 96.4%; Pred. No. 1.1;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CTCTTCTCCGAGCCGCTCGAATAAGTAGT 31
|||||
Db 22 CTATTCTCCGAGCCGCTCGAATAAGTAGT 49
|||||

RESULT 5
AX418533
LOCUS AX418533 50 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 19 from Patent WO200006.
ACCESSION AX418533
VERSION AX418533.1 GI:21523398
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.

REFERENCE 1
AUTHORS Lu, Y. and Li, J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 020006-A 19 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

FEATURES
source

Location/Qualifiers
1. .50
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Zn-DNA"

ORIGIN

Query Match 80.0%; Score 26.4; DB 6; Length 50;
Best Local Similarity 96.4%; Pred. No. 1.1;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CTCTTCTCCGAGCCGCTCGAATAAGTAGT 31
|||||
Db 22 CTATTCTCCGAGCCGCTCGAATAAGTAGT 49
|||||

```

RESULT 6
LOCUS AX418535
DEFINITION Sequence 21 from Patent WO0200006.
ACCESSION AX418535
VERSION AX418535.1 GI:215233400
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1 Lu, Y. and Li, J.
AUTHORS
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 0200006-A 21 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
source
1..50
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Zn-DNA"
ORIGIN
Query Match 72.7% Score 24; DB 6; Length 50;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 6 CTTCTCCGAGCCGGTCGGAATAGT 29
Db 6 CTTCTCCGAGCCGGTCGGAATAGT 29

RESULT 7
LOCUS AX418521
DEFINITION Sequence 7 from Patent WO0200006.
ACCESSION AX418521
VERSION AX418521.1 GI:215233386
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE
1 Lu, Y. and Li, J.
AUTHORS
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 0200006-A 7 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
source
1..33
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Variant of deoxyribozyme named 17E1"
ORIGIN
Query Match 66.1% Score 21.8; DB 6; Length 33;
Best Local Similarity 78.8%; Pred. No. 1.1e+02;
Matches 26; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Cy 1 CATCTCTTCCGAGCCGGTCGGAATAGT 33
Db 1 CATCTCTTCTCAGCGACTCGGAATAGT 33

RESULT 8
LOCUS AC112761
DEFINITION Rattus norvegicus clone CH230-51A1, *** SEQUENCING IN PROGRESS ***,
3 unordered pieces.
ACCESSION AC112761
VERSION AC112761.5 GI:25006812
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)

```

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 227565)

Murphy D.Marie, Metzker M.Lee., Abramson S., Adams C., Alder J., Allen C., Allen H., Alsbrooks S., Amin A., Anguiano D., Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H., Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed P., Bryant N., Blair O., Blankenburg K., Blyth P., Brown M., Bryant N., Buhay C., Burch P., Burrell K., Calderon E., Cardenas V., Carter K., Cavazos I., Cessier H., Center A., Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J., Cleveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L., Davila M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D., Delgado O., Denson S., Deramo C., Ding Y., Dinh H., Divya K., Draper H., Dugan-rocha S., Dunn A., Durbin K., Duval B., Eaves K., Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G., Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P., Fraser C.M., Gabis A., Ganta R., Garcia A., Garner T., Garza M., Gebregregorie E., Geer K., Gill R., Grady M., Guerra M., Guevara M., Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K., Harvey Y., Havlak P., Hawes A., Henderson N., Hernandez J., Hernandez R., Hines S., Hladun S.L., Hodgson A., Hogues M., Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A., Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Colivet A., Karpathy S., Kelly S., Kelly S., Khan Z., King L., Kovar C., Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu C., Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J., Lorenshewa L., Loulsegad H., Lozano R.J., Lu X., Ma J., Maheshwari M., Mahindaratne M., Mahmoud M., Malloy K., Mangum A., Mangun B., Mopua P., Martin K., Martin R., Martinez E., Mathew S., McLeod M.P., McNeill T.Z., Meenen E., Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S., Morgan M., Morris K., Morris S., Munitasa M., Murphy M., Nair L., Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Nwankwelen O., Okwuonu G., Olarnpunagoon A., Pal S., Parke K., Pasternak S., Paul H., Perez A., Perez L., Pfannkuch C., Plopper P., Poldester A., Popovic D., Primus E., Pu L.L., Puazo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R., Reilly B., Reilly M., Ren Y., Reuter M., Richards S., Riggs P., Rivers C., Rokey R., Rojars A., Rose M., Rose R., Ruiz S., Sanders N., Savary G., Scherer S., Scott G., Shatman S., Shen M., Shetty J., Shvartsbeyn A., Sisson I., Sitter C.D., Smajic J., Sneed A., Sodergren E., Song X.-Z., Sorelle K., Sobor P., Steidle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C., Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Usmani K., Valas R., Vera V., Villaseca D., Waldron L., Walker B., Wang J., Wang Q., Wang S., Warren J., Warren R., Wei X., White F., Williams G., Willson R., Wleczky R., Woolder H., Worley K., Wright D., Wright R., Wu J., Yakub S., Yen J., Yoon L., Yoon V., Yu P., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von Niederhausern A., Weise R., Smith D.R., Holt R.A., Smith H.O., Weinstock G. and Gibbs R.A.

Direct Submission

Unpublished

2 (bases 1 to 227565)

Worley K.C.

Direct Submission

Submitted (24-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 227565)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 15, 2002 this sequence version replaced gi:23195198.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas

COMMENT

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the features table.

----- Genome Center
Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: INFO

Center clone name: CH230-51A1

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 216274 bases at least Q40

Consensus quality: 218805 bases at least Q30

Consensus quality: 220176 bases at least Q20

Estimated insert size: 223866; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 222502: contig of 22502 bp in length

* 222503 222502: gap of unknown length

* 222603 222602: contig of 1028 bp in length

* 223631 223730: gap of unknown length

* 223731 227565: contig of 3835 bp in length.

* Location/Qualifiers

1. 227565

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-51A1"

1. 1192

/note="wgs end extension"

clone_end:77"

4537. 5317

/note="wgs boundary"

clone_end:77

site:

end sequence: BH336817"

114574. 119241

/note="wgs contig"

163419. 165095

/note="wgs contig"

ORIGIN

Query Match 64.8% Score 21.4; DB 2; Length 227565;

Best Local Similarity 80.6% Pred. No. 3.2e+02;

Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TCTCTCTCCAGCCCGCTCCAAATAGTGAGT 33

DB 80607 TCTCTCTCTCCAGCCCGCTCCAAATAGTGAGT 80637

RESULT 9

AC118309

LOCUS

DEFINITION

AC118309 234801 bp DNA linear HTG 13-NOV-2002

Rattus norvegicus clone CH230-212018. *** SEQUENCING IN PROGRESS

***, 5 unordered pieces.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AC118309

AC118309.6 GI:24941765

HTG, HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 234801)

Murphy, D.Marie., Metzker, M.Lee., Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, K., Alabrooks, S., Amin, A., Angulano, L.,

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Banderamike, D., Barber, M., Barnstead, K., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyie, M., Cree, A., D'Souza, L.,

Davila, M.J., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Secotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregorgis, E., Gear, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,

Gumaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Hartney, Y., Havlak, P., Hayes, A., Henderson, N., Hernandez, J.,

Herrera, R., Hines, S., Hladun, S.J., Hodgson, A., Hughes, M.,

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idelbird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, R., Johnson, B., Johnson, R., Jollivet, A.,

Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, D., Kovar, C.,

Kiwi, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorenz, L., Loulseg, H., Lozano, R.J., Lu, X., Ma, J.,

Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Margum, A.,

Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,

Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Narf, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwokedin, O., Obedun, G., Ocarin, S., Ocarin, S., Ocarin, S.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Frankoch, C.,

Plopper, P., Poldinger, A., Popovic, D., Primus, E., Pu, L.-L.,

Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,

Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P.,

Rives, C., Rodkey, I., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,

Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,

Shetty, J., Shvartabeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,

Sheele, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,

Stefan, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,

Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejo, Z., Umani, K.,

Valas, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J.,

Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P.,

Williams, G., Willson, R., Wleciyk, R., Woodson, H., Worley, K.,

Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,

Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von

Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.C.,

Weinstock, G. and Gibbs, R.A.

Direct Submission

Unpublished

2 (bases 1 to 234801)

Worley, K.C.

Direct Submission

Submitted (15-APR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 234801)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (13-NOV-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Nov 13, 2002 this sequence version replaced gi:23267374.

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genomes Center
Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GUTA

Center clone name: CH230-212018

----- Summary Statistics

Assembly program: Phrap; version 0.90329

Consensus quality: 227577 bases at least Q40

Consensus quality: 230382 bases at least Q30

Consensus quality: 231817 bases at least Q20

Estimated insert size: 239036; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 162694: contig of 162694 bp in length
* 162695 162794: gap of unknown length
* 162795 225625: contig of 52831 bp in length
* 225626 235725: gap of unknown length
* 235726 231241: contig of 5516 bp in length
* 231242 231341: gap of unknown length
* 231342 232640: contig of 1299 bp in length
* 232641 232740: gap of unknown length
* 232741 234801: contig of 2061 bp in length.

FEATURES

source
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-212018"
complement(160419..161185)
/note="clone_boundary
clone_end:T7
site:
end sequence:BNH264302"
225726..227372
/note="wgs end_extension
clone_end:T7"

misc_feature

misc_feature

ORIGIN

Query Match 64.8%; Score 21.4; DB 2; Length 234801;
Best Local Similarity 80.6%; Pred. NO. 3.2e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TCTCTTCGAGCCCGTCGAAATAGTGAGT 33
|||||
DB 222506 TCTCTTCGAGCCCATCCAAATAGTAGT 222536
|||||

RESULT 1C

AC094031

LOCUS

DEFINITION

AC094031

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 244180)

Murphy, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, P., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Arca, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Ditz, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Garcia, R., Gant, A., Garner, T., Garza, M., Gbagrege, B., Geet, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Guarnatone, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, J., Harvey, I., Havlik, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenz, L., Louie, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, N., Mahmoud, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawlawi, S., McLeod, M. P., McNeil, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minje, S., Muniyasa, M., Murphy, M., Naif, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, P., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R., Rivas, C., Rodkey, T., Royce, A., Rose, M., Rowe, K., Ruiz, S. J., Sanders, N., Savory, G., Scherer, S., Scott, G., Shatsman, S., Shan, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajda, D., Shedd, J., Sodergren, B., Song, X. Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, R., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlezyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yan, J., Yoon, V., Yu, P., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

2 (bases 1 to 244180)

Worley, K. C.

Direct Submission

Unpublished

2 (bases 1 to 244180)

Worley, K. C.

Direct Submission

Submitted (14-SEP-2001)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 244180)

Rat Genome Sequencing Consortium.

Direct Submission

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On May 10, 2003 this sequence version replaced gi:24942156. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRQA
Center clone name: CH230-53K20
----- Summary Statistics
Assembly program: Atlas 3.0
Consensus quality: 209969 bases at least Q40
Consensus quality: 215543 bases at least Q30
Consensus quality: 218747 bases at least Q20
Estimated insert size: 220505, sum-of-contigs estimation
Quality coverage: 6x in Q20 bases, sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 241874: contig of 241874 bp in length
2 241875: gap of unknown length
3 241876: contig of 1035 bp in length
4 243010: gap of unknown length
5 243110: contig of 1071 bp in length.

FEATURES

source
1..244180
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-53K20"
misc_feature
1..2594
/note="wgs_contig"
misc_feature
101472..103123
/note="wgs_contig"
misc_feature
113845..115536
/note="wgs_contig"
misc_feature
236001..237221
/note="wgs_contig"

ORIGIN

Query Match 64.8%; Score 21.4; DB 2; Length 244180;
Best Local Similarity 80.6%; Pred. No. 3.2e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TCTCTCTCCGAGCCCGCTCGAATAGTAGT 33
DB 238824 TCTCTCTCCGAGCCCGCTCGAATAGTAGT 238854

RESULT 11

SCU2022a/c

LOCUS
DEFINITION

3356 bp DNA linear PLN 14-JUL-1995
Saccharomyces cerevisiae RNA polymerase II holoenzyme
cyclic-dependent kinase component catalytic subunit (SRB10) gene,
complete cds.

ACCESSION

U20222

U20222.1

GI:662799

Saccharomyces cerevisiae (baker's yeast)

Saccharomyces cerevisiae

Rukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

1 (bases 1 to 3356)

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

Li, S.M., Zhang, J., Jeffery, D.A., Koleske, A.J., Thompson, C.M.,

Chao, D.M., Viljoen, M., van Vuuren, H.J., and Young, R.A.

A kinase-cyclin pair in the RNA polymerase II holoenzyme

Nature 374 (6518), 193-196 (1995)

95183147

7877695

2 (bases 1 to 3356)

Li, S.

Direct Submission

Submitted (24-JAN-1995) Sha-Mei Liao, Whitehead Institute for

Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142,

USA

Location/Qualifiers

1..3356

/organism="Saccharomyces cerevisiae"

/mol_type="genomic DNA"

/strain="S288C"

/db_xref="taxon:4932"

/chromosome="XVI"

/map="next to ISM1, see GenBank Accession Number L38957"

1302..2968

/gene="SRB10"

1302..2968

/gene="SRB10"

1303..2968

/gene="SRB10"

1343..2968

/gene="SRB10"

/functions="influences the phosphorylation of the largest

subunit of RNA polymerase II (CTD)"

/note="catalytic subunit of a cyclin-dependent kinase

component of RNA polymerase II holoenzyme; required for

responses to positive and negative regulators; gene also

called UME5, GenBank Accession Number L27151"

/codon_start=1

/product="Srb10p"

/protein_id="AAC13785.1"

/db_xref="GI:662800"

/translation="MYCRPMQVCCQQAQSVKQKQIGSVHGKAPMLMANNVFTIG

PYARKRMEVSVLEKRVIGVYIAAGTYGKVKYKAGQINCTGANGSSNGTNAKIP

QFDTQPKSSSDMDQNTALRNLLKDEGVTGCRITREDVSHPNSQKQTLIKK

PLTVFAIKKFTKEDVEQLHYTGISQACREMLCRLEINLITLTVIPERKCY

HMVYASDILLQIHFHSHPEKMIIPPMVRSIMQLDGVSLHOMVLRDLKPA

NIMVYIDGCVKIGDLGKARFNNMLTLYTGDVQVVTWYRAPELLGARYTPVDL

WSVCCIPAKLIGLPFKGHEAKLDSKTVPPFVNQLQRIEVLGTPDKIMPYLEKY

PEYDQITKPKYRDNLATWYHAGGGRDKSLYLHLLNVDPIKRIDAFVALLKPYT

RSIDIPVSEVPEGLTYKYPAREHTNDIMVLSRTKNTNTAGSITAGAAANUGSL

GVNRRILLAAAAAANAVSGNADSPSKKRR"

ORIGIN

Query Match 63.6%; Score 21; DB 8; Length 3356;
Best Local Similarity 82.8%; Pred. No. 3.5e+02;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 4 CTCCTCTCCGAGCCCGCTCGAATAGTAGT 32
DB 537 CGTGTCTCTGAGCGGCTCGAATAGTAGT 509

RESULT 12

```

AX194850
LOCUS AX194850 475 bp DNA linear PAT 28-AUG-2001
DEFINITION Sequence 320 from Patent WO0151659.
ACCESSION AX194850
VERSION AX194850.1 GI:15385497
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Chu, T., Blumenfeld, M. and Cohen, D.
TITLE Biallelic markers derived from genomic regions carrying genes
involved in central nervous system disorders
JOURNAL Patent: WO 0151659-A 320 19-JUL-2001;
GENSET (FR)
FEATURES
source
1..475
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
primer_bind
1..20
/note="upstream amplification primer"
misc_feature
332..351
/note="18-523-352.mis1, potential"
misc_feature
340..364
/note="18-523-352 potential probe"
variation
352
/note="18-523-352 : polymorphic base A or G"
misc_feature
353..371
/note="18-523-352.mis2, complement"
primer_bind
455..475
/note="downstream amplification primer, complement"
ORIGIN
Query Match 62.4%; Score 20.6; DB 6; Length 475;
Best Local Similarity 85.2%; Pred. No. 4.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 TCTCTCTCCGAGCCGGTTCGAATAGT 29
DB 11 TCTCTCTCCGAGCCGGTTCGAATAGT 37

RESULT 13
AX467748 1248 bp DNA linear PAT 16-JUL-2002
LOCUS AX467748
DEFINITION Sequence 12 from Patent WO0236620.
ACCESSION AX467748
VERSION AX467748.1 GI:21900918
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Rueegg, M.A. and Brugnot, A.
TITLE Ets-transcription factor related compound specific promoter and
transactivators thereof
JOURNAL Patent: WO 0236620-A 12 10-MAY-2002;
Myocontract Pharmaceutical Research AG (CH)
FEATURES
Location/Qualifiers
source
1..1248
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="DNA sequence encoding GABPbeta_VP16"
ORIGIN
Query Match 62.4%; Score 20.6; DB 6; Length 1248;
Best Local Similarity 85.2%; Pred. No. 4.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 TCTTCTCCGAGCCGGTTCGAATAGTCA 31

```

```

DB 168 TCTTCTCCGAGCCGGTTCGAATAGTCA 194

RESULT 14
MUSGAC 1338 bp mRNA linear ROD 09-NOV-1994
LOCUS MUSGAC
DEFINITION Mouse GA binding protein (GABP-beta2 subunit) mRNA, complete cds.
ACCESSION M74517 M74517
VERSION M74517.1 GI:193412
KEYWORDS GA-binding protein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1338)
AUTHORS LaMarco, K., Thompson, C.C., Byers, B.P., Walton, E.M. and
McKnight, S.L.
TITLE Identification of Ets- and notch-related subunits in GA binding
protein
JOURNAL Science 253 (5021), 789-792 (1991)
MEDLINE 91343912
PubMed 1876936
COMMENT Original source text: Mus musculus (strain CD-1) 8.5 day old embryo
cDNA to mRNA.
FEATURES
Location/Qualifiers
source
1..1338
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CD-1"
/db_xref="taxon:10090"
/dev_stage="8.5 day old embryo"
gene
1..1338
/gene="GABP"
CD8
132..1178
/gene="GABP"
/note="ORP"
/codon_start=1
/product="GA binding protein"
/protein_id="AA53032.1"
/db_xref="GI:567203"
/translation="MSLVDSGKLEAPAGODDEVRLIMANGAPFTTGLSTPLHL
AAQGHFTTEVLLAGVSDATKYDRFLPMAAGEGNAIVEVLKGGADVAKOM
LKTALHWAHGHGVEVLLIKYGDVHTQKCTKAFDLSISNGREDLARIQIAM
QKQINTPEPDPVTIHAATPQFIQGGVNLDTDETGVSAVPGNSS-SVALTIAL
AASAPLNSSETFVVAIEHVTAESVDGAIQQVSSGGQVITIVTGGIQGNLH1
PTSGGQPIIVTTPDQQQVLTVPATDIABETVISPEPPAKKQCMELIESRVECAIEV
ESLIPGVFCCSHPK"
ORIGIN
Query Match 62.4%; Score 20.6; DB 10; Length 1338;
Best Local Similarity 85.2%; Pred. No. 4.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 5 TCTTCTCCGAGCCGGTTCGAATAGTCA 31
DB 299 TCTTCTCCGAGCCGGTTCGAATAGTCA 325

RESULT 15
MUSGAB 2598 bp mRNA linear ROD 09-NOV-1994
LOCUS MUSGAB
DEFINITION Mouse GA binding protein (GABP-beta1 subunit) mRNA, complete cds.
ACCESSION M74516 M74516
VERSION M74516.1 GI:193384
KEYWORDS GA-binding protein.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2598)
AUTHORS LaMarco, K., Thompson, C.C., Byers, B.P., Walton, E.M. and
McKnight, S.L.

```


TITLE Identification of Ets- and notch-related subunits in GA binding protein
 JOURNAL Science 253 (5021), 789-792 (1991)
 MEDLINE 91343912
 PUBMED 1876836
 COMMENT Original source text: Mus musculus (strain CD-1) 8.5 day old embryo cDNA to mRNA.

FEATURES
 source
 1..2598
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="CD-1"
 /db_xref="taxon:10090"
 /dev_stage="8.5 day old embryo"
 1..2598
 /gene="GABP"
 133..1281
 /gene="GABP"
 /note="beta1 subunit"
 /codon_start=1
 /product="GA binding protein"
 /protein_id="AA053031.1"
 /db_xref="GI:567202"
 /translation="MSLVGLGKLLKLEAARAGQDDEVRIILMANGAPPTDWTGTSPIHL
 AAGYGFSTVTEVLLAGVSRDARTKVTPLHMAEGHANTVEVLLKGGADVNAQIM
 LKMTALRNATENHQQVVELLIKYGADVTKQKCTAPDIDNGNEDLAEIQLAM
 QNQTNPSPDTVTHAATPQFIIGPGVNLDTGTVGAVPGNSSTVLATLAL
 ABASPLNSSTPTVATREVTASVDGAIQVSSGGQVITVTDGILGNLHSI
 TSGMGQPIITVTPDQGVLTVPATDIAETVISEEPPAKRQCKHIESRVECAETEE
 EALQQLDANRBAQYKQQLKKEQBABAYRQKLEAMTRCTQNKBAV"

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 2598;
 Best Local Similarity 85.2%; Pred. No. 5.1e-02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TCTTCTCCGAGCCGGTCGAATAGTGA 31
 |||||
 Db 300 TCTTCTCCGAGCCGGTGAAGTAGAGA 326

RESULT 16
 LOCUS BC013558 2622 bp mRNA linear ROD 01-OCT-2003
 DEFINITION Mus musculus GA repeat binding protein, beta 1, mRNA (cDNA clone
 MGC:6203 IMAGE:3592119), complete cds.
 BC013558
 VERSION BC013558.1 GI:15488854
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2622)
 Strausberg R.L., Feingold R.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh Y.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,
 Scheetz T.E., Brownstein M.J., Uudin T.B., Toshiyuki S.,
 Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
 Abramson R.D., Mullenbach S.J., Bosak S.A., McEwan P.J.,
 McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
 Varley K.C., Hale S., Garcia A.M., Gay D.J., Hulyk G.B.,
 Villalón D.K., Mundy P.M., Sodergren S.J., Lu X., Gibbs R.A.,
 Paney J., Heiton S., Kettman M., Madan A., Rodriguez S.,
 Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y.,
 Bouckard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
 Dickson M.C., Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,
 Butterfield V.S., Krzywinski M.I., Skalek U., Smallos D.B.,
 Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
 Generation and initial analysis of more than 15,000 full-length

TITLE

JOURNAL human and mouse cDNA sequences
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 22388257
 12477932
 REFERENCE 2 (bases 1 to 2622)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 Submitted (04-SEP-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: ggapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The J.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: http://www-shgc.stanford.edu
 Contact: (Dickson, Mark) mcdpaxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the J.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAC Plate: 11 Row: 4 Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA GI: 5753933.
 Location/Qualifiers

FEATURES
 source

1..2622
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="CZBCH II"
 /db_xref="taxon:10090"
 /clone="MGC:6203 IMAGE:3592119"
 /tissue_type="Mammary tumor metastasized to lung; Tumor
 arose spontaneously from a senescent normal mammary
 (clonal) outgrowth infected with the virus MMTV."
 /clone_lib="NCI CCAP_L029"
 /lab_host="DHI08"
 /note="Vector; PCMV-SPORT6"
 1..2622
 /gene="Gabp1"
 /db_xref="LocustID:14391"
 /db_xref="MGI:95611"
 129..1280
 /codon_start=1
 /product="Gabp1 protein"
 /protein_id="AAH13558.1"
 /db_xref="GI:15488855"
 /db_xref="LocustID:14391"
 /translation="MSLVGLGKLLKLEAARAGQDDEVRIILMANGAPPTDWTGTSPIHL
 AAGYGFSTVTEVLLAGVSRDARTKVTPLHMAEGHANTVEVLLKGGADVNAQIM
 LKMTALRNATENHQQVVELLIKYGADVTKQKCTAPDIDNGNEDLAEIQLAM
 QNQTNPSPDTVTHAATPQFIIGPGVNLDTGTVGAVPGNSSTVLATLAL
 ABASPLNSSTPTVATREVTASVDGAIQVSSGGQVITVTDGILGNLHSI
 TSGMGQPIITVTPDQGVLTVPATDIAETVISEEPPAKRQCKHIESRVECAETEE
 EALQQLDANRBAQYKQQLKKEQBABAYRQKLEAMTRCTQNKBAV"
 141..530
 /note="ATP; Region: POG: Acyrin repeat [General function
 prediction only]"
 /db_xref="CDD:COG0666"

gene

CDS

misc_feature

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 2622;
 Best Local Similarity 85.2%; Pred. No. 5.1e-02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Qy 5 TCTTCTCCGAGCCGGTCGAATAGTGA 31
 |||||
 Db 296 TCTTCTCCGAGCCGGTGAAGTAGAGA 322

RESULT 17
 AL844555/c
 LOCUS
 DEFINITION
 Mouse DNA sequence from clone RP23-251B2 on chromosome 2, complete sequence.
 ACCESSION
 AL844555
 VERSION
 HTG.
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 12221)
 Leongamornlert.D.
 Direct Submission
 Submitted (15-NOV-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Nov 19, 2002 this sequence version replaced g1:2493948.
 Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Web site: <http://www.sanger.ac.uk>
 Contact: humquerry@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one N13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sv, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-251B2 is from the RPCT-23 Mouse PAC Library
 Constructed by the group of Pieter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBAC3.6.

FEATURES

Location/Qualifiers
 1..12221
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-251B2"
 /clone_lib="RPCT-23"

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 12221;
 Best Local Similarity 85.2%; Pred. No. 6.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGACCGCGTGCMAATAGTGA 31
 |||||
 Db 82878 TCTTCTCCGACCGCGTGTACTAGAGA 82852
 |||||

RESULT 18

AC008684/c
 LOCUS
 DEFINITION
 Homo sapiens chromosome 5 clone CTB-58H17, WORKING DRAFT SEQUENCE, 21 ordered pieces.
 ACCESSION
 AC008684
 VERSION
 HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEPIN.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 141132)
 DOE Joint Genome Institute.
 Sequencing of Human Chromosome 5
 Unpublished
 2 (bases 1 to 141132)
 DOE Joint Genome Institute.
 Direct Submission
 Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Jul 18, 2000 this sequence version replaced g1:7709264.
 ----- Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.jgi.doe.gov>

Summary Statistics

Consensus quality: 113678 bases at least Q40
 Consensus quality: 129742 bases at least Q30
 Consensus quality: 135169 bases at least Q20
 Estimated insert size: 153000; pulse field gel estimation
 Estimated insert size: 140132; sum-of-contigs estimation
 Quality coverage: 4.31 in Q20 bases; pulse field gel estimation
 Quality coverage: 4.71 in Q20 bases; sum-of-contigs estimation.
 NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
 * This sequence will be replaced
 * By the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 15398: contig of 15398 bp in length
 * 15499: gap of unknown length
 * 37419: contig of 21921 bp in length
 * 37420: gap of unknown length
 * 37519: gap of unknown length
 * 44824: contig of 7305 bp in length
 * 44825: gap of unknown length
 * 44925: gap of unknown length
 * 54089: gap of unknown length
 * 54189: gap of unknown length
 * 57422: contig of 3234 bp in length
 * 57423: gap of unknown length
 * 57523: gap of unknown length
 * 65914: gap of unknown length
 * 66015: gap of unknown length
 * 68688: contig of 2674 bp in length
 * 68789: gap of unknown length
 * 70851: contig of 2063 bp in length
 * 70852: gap of unknown length
 * 70952: gap of unknown length
 * 86942: contig of 15991 bp in length
 * 86943: gap of unknown length
 * 87043: gap of unknown length
 * 89989: contig of 7947 bp in length
 * 89990: gap of unknown length
 * 90090: gap of unknown length
 * 94658: contig of 4569 bp in length
 * 94758: gap of unknown length
 * 94759: gap of unknown length
 * 99472: contig of 4714 bp in length
 * 99572: gap of unknown length
 * 99573: gap of unknown length
 * 102375: contig of 2703 bp in length
 * 102376: gap of unknown length

Tue May 25 08:19:54 2004

```

* 102376 104007: contig of 1632 bp in length
* 104008 104107: gap of unknown length
* 105108 105117: contig of 1810 bp in length
* 105918 106017: gap of unknown length
* 106018 106218: contig of 3911 bp in length
* 109929 110028: gap of unknown length
* 110029 110945: contig of 6917 bp in length
* 110946 117045: gap of unknown length
* 117046 122158: contig of 5113 bp in length
* 122159 122515: gap of unknown length
* 122516 125715: gap of unknown length
* 125716 130631: contig of 4915 bp in length
* 130631 130731: gap of unknown length
* 130731 131132: contig of 10402 bp in length.
Location/Qualifiers
1. .141132
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="5"
/clone="CTB-58H17"
/clone_lib="CalTech human BAC library B"

ORIGIN
Query Match 62.4%; Score 20.6; DB 2; Length 141132;
Best Local Similarity 85.2%; Pred. No. 6.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TCTCTTCTCCGAGCCGTCGAAATAGT 29
|||||
DB 45114 TCTCGTCAACGAGCCGTCGAAATAGT 45088

RESULT 19
AP165124 195909 bp DNA linear PRI 07-JUN-2001
LOCUS Homo sapiens chromosome 5q31.1-q33.1 clone BAC dln082c10 containing
DEFINITION GABRG2 gene, complete sequence.
ACCESSION AP165124
VERSION AP165124.1 GI:5738137
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 195909)
AUTHORS Jiang,S., Yu,J., Wang,J., Tan,Z., Xue,H., Peng,G., He,L. and
Yang,H.
TITLE Complete genomic sequence of 195 Kb of human DNA containing the
gene GABRG2
JOURNAL DNA Seq. 11 (5), 373-382 (2000)
MEDLINE 21226276
PubMed 11328646
REFERENCE 2 (bases 1 to 195909)
AUTHORS Jiang,S., Yu,J., He,L. and Yang,H.
TITLE Direct Submission
JOURNAL Submitted (02-JUL-1999) Human Genome Center, Institute of Genetics,
Chinese Academy of Sciences, Datun Road, Beijing 100101, People's
Republic of China
Location/Qualifiers
1. .195909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="5"
/map="5q31.1-q33.1"
/clone="BAC dln082c10"
/complement("<70096..>155468)
/gene="GABRG2"

gene
complement(join("<70096..70371,74151..74356,81148..81300,
119446..119583,122155..122237,125614..125834,
127910..127977,129493..129644,155362..155468))

FEATURES
source
1. .195909
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="5"
/map="5q31.1-q33.1"
/clone="BAC dln082c10"
/complement("<70096..>155468)
/gene="GABRG2"

CDS
/genes="GABRG2"
/product="gamma-aminobutyric acid A receptor gamma 2"
complement(join(70096..70371,74151..74356,81148..81300,
119446..119583,122155..122237,125614..125834,
127910..127977,129493..129644,155362..155468))
/gene="GABRG2"
/codon_start=1
/product="gamma-aminobutyric acid A receptor gamma 2"
/protein_id="AA50273.1"
/db_xref="GI:5738138"
/translation="MSSPNWTSQSSVYSPVQRTWILLLSYRGFTSQKDD
DYEDYASNTWLTTPKVEGDTVLNNLLEGYDNLKRPDQIGVKPLIHTDMVNSIG
PVNINWYTIIDIFPAOTYDRRLKFNSTIKVLALNSNMVGRKWIPTDFRNGKADA
HNITTPRMLRWNGRVLVLTILRTIDAEQLQHNFPMDHSCLFESGYPREI
VYQWRSSVEVDGTRSLYQSPVGLRNTTEVVKTTSGDYVVMVYFDLSRMGYFT
IQTYIPTLIVLSHW8PWINKDAVPARTSLGITVLTTLSTIARKSLPKVSYVTA
MDLPVSVCFIPVFSALVRYGLRYPVGRKFP8DKDKCKKNPAPTIDIRPSATIQN
NATHLQRDEEVGYECLDQDCNSFPFCFBCDRTGAWRHGRIHIRIAKDCSYARIFPP
TATCUFNUVWYSYLYL"

ORIGIN
Query Match 62.4%; Score 20.6; DB 9; Length 195909;
Best Local Similarity 85.2%; Pred. No. 7.1e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TCTCTTCTCCGAGCCGTCGAAATAGT 29
|||||
DB 117053 TCTCGTCAACGAGCCGTCGAAATAGT 117079

RESULT 20
AC091926 209836 bp DNA linear PRI 21-SEP-2001
LOCUS Homo sapiens chromosome 5 clone RP11-290G6, complete sequence.
DEFINITION AC091926
ACCESSION AC091926
VERSION AC091926.4 GI:15718547
KEYWORDS HTG.
SOURCE DOB Joint Genome Institute and Stanford Human Genome Center.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 209836)
AUTHORS DOB Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 2 (bases 1 to 209836)
AUTHORS DOB Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 209836)
AUTHORS DOB Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (30-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE 4 (bases 1 to 209836)
AUTHORS DOB Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Sep 21, 2001 this sequence version replaced gi:15375175.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.3.
STS Content:
SHGC-79397 652187.
Location/Qualifiers
1. .209836
/organism="Homo sapiens"
/mol_type="genomic DNA"

```

Tue May 25 08:19:54 2004

Sloan-Kettering Cancer Center, 1275 York Ave. RRL 937, New York, NY
10021, USA

FEATURES

source

Location/Qualifiers
1. .17683
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="RII1"
/db_xref="taxon:10090"
/chromosome="12"
/map="55 CM"
/clone_lib="Genome Systems PI RII1 genomic library"
/notes="partial sequence of a 26 kb clone"

misc_feature

1. .17683
/note="contains 3' end of MRCKbeta gene, similar to 3' end of rat MRCKbeta and Homo sapiens CDC42BP genes"

ORIGIN

Query Match 61.8%; Score 20.4; DB 10; Length 17683;
Best Local Similarity 80.0%; Pred. No. 7.2e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 TCTCTCTCCGAGCGCGTGGAAATAGT 32

DB 16564 TCTCTCTCCGAGCGCGTGGAAATAGT 16535

RESULT 23

AC109717

LOCUS

DEFINITION

AC109717

AC109717.4 GI:23820695

VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.

KEYWORDS

Rattus norvegicus (Norway rat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE

AUTHORS

1 (bases 1 to 226911)

Munoz, D. Marie, Metzger, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Albrechts, S., Amin, A., Anguiano, D.,

Anyalebechi, V., Ayegba, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Devilla, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dedrich, D.,

Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Evans, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flegg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregorgis, S., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, M.,

Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,

Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, J.,

Hollins, B., Howells, S., Huly, S., Hume, J., Idlebird, D., Jackson, A.,

Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpach, S., Kelly, S., Khan, Z., King, L., Kovar, C.,

Kowit, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,

Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,

Lorenshew, L., Louissegh, M., Lozano, R. J., Lu, X., Ma, J.,

Maheshwari, M., Mahindratne, M., Mahmood, M., Mallory, K., Mangum, A.,

Mangum, B., Mapus, P., Martin, X., McNeill, T. Z., Meenen, E.,

Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,

Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,

Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,

Nwokolemah, O., Okwono, G., Olarnpunsagoon, A., Pal, S., Parks, K.,

Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,

/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-290G6"

ORIGIN

Query Match 62.4%; Score 20.6; DB 9; Length 209836;
Best Local Similarity 85.2%; Pred. No. 7.1e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 TCTCTCTCCGAGCGCGTGGAAATAGT 29

DB 180805 TCTCTCTCCGAGCGCGTGGAAATAGT 180831

RESULT 21

AX658830

LOCUS

DEFINITION

AX658830

AC109717.4 GI:23820695

VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.

KEYWORDS

Rattus norvegicus (bread wheat)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Poideae; Triticeae; Triticum.

REFERENCE

AUTHORS

Sessions, A., Briggs, S., Cooper, B., Goff, S. A., Moughamer, T., and

Glazebrook, J., Katagiri, F., Kreps, J., Provart, N., Riecke, D., and

Zhu, T.

TITLE

Identification and characterization of plant genes

JOURNAL

Patent: WO 0300904-A 103 03-JAN-2003;

Syngenta Participations AG (CH)

FEATURES

source

1. .1254

/organism="Triticum aestivum"

/mol_type="unassigned DNA"

/db_xref="taxon:4565"

ORIGIN

Query Match

Best Local Similarity

Matches

24; Conservative

0; Mismatches

6; Indels

0; Gaps

0;

QY

1 CATCTCTCCGAGCGCGTGGAAATAGT 30

DB

738 CTCTCTCTCCGAGCGCGTGGAAATAGT 767

RESULT 22

AF320617/c

LOCUS

DEFINITION

AF320617

AC109717.4 GI:23820695

VERSION

HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.

KEYWORDS

Rattus norvegicus (house mouse)

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 17683)

Kalanthy, S., Manning, S., Haub, O., Tomihara-Newberger, C., Lee, H. G.,

Pangman, J., Dische, C. M., Manova, K. and Lacy, E.

The amnionless gene, essential for mouse gastrulation, encodes a

visceral-endothelial-specific protein with an extracellular

cysteine-rich domain

Nat. Genet. 27 (4), 412-416 (2001)

21175746

JOURNAL

MEDLINE

PUBMED

11279523

REFERENCE

2 (bases 1 to 17683)

Kalanthy, S., Haub, O. and Lacy, E.

Direct Submission

TITLE

Submitted (13-NOV-2000) Molecular Biology Program, Memorial

Plopper, P., Poindestre, A., Popovic, D., Primus, R., Pu, L.-L., Puazo, M., Quirio, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Kelly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojase, A., Rose, M., Rose, R., Ruiz, S., Sanders, M., Savary, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steale, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Wang, S., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wleczek, R., Wood, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Yen, J., Yoon, L., Yoon, V., Niederhauern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 226911)
Worley, K.C.

Submitted (07-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 226911)

Rat Genome Sequencing Consortium.
Direct Submission
Submitted (11-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Oct 11, 2002 this sequence version replaced gi:21738139.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.bgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the contigs and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.bgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: Q2K
Center clone name: CH230-302F7
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 215934 bases at least Q40
Consensus quality: 215943 bases at least Q30
Consensus quality: 216975 bases at least Q20
Estimated insert size: 216557; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.bgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved
* 1 225696: contig of 225696 bp in length
* 225697 225796: gap of unknown length
* 225797 226911: contig of 1115 bp in length.
* Location/Qualifiers

source
1. 226911
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-302F7"
1. 1189
/note="wgs end extension
clone_end:Sp6"
1240. 2418
/note="wgs end extension
clone_end:Sp6"
6786. 7661
/note="clone boundary
clone_end:Sp6"
site:RboI
end sequence:EXAKG28TV"
217599. 218035
/note="clone boundary
clone_end:T7
site:RboI
end sequence:EXAKG28TV"
ORIGIN
Query Match 61.8%; Score 20.4; DB 2; Length 226911;
Best Local Similarity 80.0%; Pred. No. 8.8e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 4 CTCTTCGCGCGCGTCCGATATAGTCAGT 33
DB 85108 CTCTTACCCAGCTGCTCTTACAGTCAGT 85137
RESULT 24
MLEPRTN3/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
FORMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE

MLEPRTN3 312050 bp DNA linear BCT 20-FEB-2001
Mycobacterium leprae strain TN complete genome; segment 3/10.
AL583919 AL450380
AL583919.1 GI:13092766
Mycobacterium leprae
Mycobacterium leprae
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
1 (bases 1 to 312050)
Cole, S.T., Eigmeier, K., Parkhill, J., James, K.D., Thomson, N.R.,
Wheeler, P.R., Honore, N., Ganier, T., Churcher, C., Harris, D.,
Mungall, K., Basham, D., Brown, D., Chillingworth, T., Connor, R.,
Davies, R.M., Devlin, K., Duthoy, S., Feltwell, T., Fraser, A.,
Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Lacroix, C.,
Maclean, J., Moule, S., Murphy, L., Oliver, Quail, M.A.,
Rajandream, M.-A., Rutherford, K.M., Rutter, S., Seeger, K., Simon, S.,
Simmonds, M., Skelton, J., Squares, R., Squares, S., Stevens, K.,
Taylor, K., Whitehead, S., Woodward, J.R. and Barrell, B.G.
Massive gene decay in the leprosy bacillus
Nature 405 (6823), 1007-1011 (2001)
2128732
11234002
2 (bases 1 to 312050)
Parkhill, J.
Direct Submission
Submitted (20-FEB-2001) Submitted on behalf of the Mycobacterium
leprae sequencing team, The Sanger Centre, Wellcome Trust Genome
Campus, Hinxton, Cambridge, CB10 1SA, UK Unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724, Paris Cedex, France. E-mail: parkhill@sanger.ac.uk
Notes:
Details of M. leprae sequencing at the Sanger Centre are available
from http://www.sanger.ac.uk/projects/M_leprae/ A relational
database containing the M. leprae sequences is available from
<http://genolist.pasteur.fr/leproma/>.
Location/Qualifiers
1. 312050
/organism="Mycobacterium leprae"

```

/mol_type="genomic DNA"
/strain="TN"
/db_xref="taxon:1769"
complement(2412..3665)
/note="LEPRT, copy 1, 1254 bp, shows 99-100% identity with regions 72707..728460, 1539937..1541188, 1480312..1481018, 2203983..2204515, 691673..691731"
complement(2413..2473)
/note="61 bp repeat found at 3'-end of LEPRT and 5' to REPLP"
complement(5611..6426)
/gene="ML0569"
complement(5611..6426)
/gene="ML0569"
/note="Similar to M. tuberculosis Rv1433, putative membrane protein, TR:006825 (EMBL:AL123456) (271 aa); Pista score E(): 0, 68.3% identity in 271 aa overlap. Shows weak similarity to other hypothetical proteins in M. tuberculosis. Previously sequenced as TR:049706 (EMBL:U00013) (271 aa); Pista score E(): 0, 100.0% identity in 271 aa overlap. Similar to ML0426, ML2446 and ML2664"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAC30077.1"
/db_xref="GI:13092767"
/db_xref="SPT:EMBL:Q49706"
/translation="MGIMRAIRGVVVLVIGIANVTIAGPVTVSLAAVNHGPAITSI LPTEDAVGVAVPVVTVTGNPTVNRKATLSLKYRSAPANRPFOLDSKVQVWPR YFEETALVITGNLSTNFTGPAIGIADINSHTFTVTDGVEATPTTGPSPWPR HXGHEWPAWSMGTEPPTPTGKTVMSKDRSVIMSSVGIPTDDEGRLVYDAV RISSRLYTHSPMAVQSMGPDNTSHGICISLPADAFTYNTYNGDPVYKX"
6876..6882
/note="possible RBS"
6889..7908
/gene="gap"
/note="synonym: ML0570"
6889..7908
/gene="gap"
/EC_number="1.2.1.12"
/note="Similar to M. tuberculosis gap, Rv1436, glyceraldehyde 3-phosphate dehydrogenase, SW:G3P_MYCTU (006822) (339 aa); Pista score E(): 0, 89.1% identity in 339 aa overlap. Similar to many e.g. Mycobacterium avium gap, glyceraldehyde 3-phosphate dehydrogenase, SW:G3P_MYCAV (P94915) (339 aa); Pista score E(): 0, 87.9% identity in 339 aa overlap. Previously sequenced as SW:G3P_MYCLE (P46713) (339 aa); Pista score E(): 0, 99.7% identity in 339 aa overlap. Contains pfam match to entry PF00044 gpdh, glyceraldehyde 3-phosphate dehydrogenase. Contains PF00071 Glyceraldehyde 3-phosphate dehydrogenase active site."
/codon_start=1
/transl_table=11
/product="glyceraldehyde 3-phosphate dehydrogenase"
/protein_id="CAC30078.1"
/db_xref="GI:13092768"
/db_xref="GOA:P46713"
/db_xref="SWISS-PROT:P46713"
/translation="MTVRVINGFGRIQRNFTYALLAQSGHGIDVQVVAINDITNS TLAYLLFDSILGRLPDVSLEEDTIVGSEKIALAVREGPAALPMAVGVVVVE STGLTNAKAGHLEAGAKVIVSAPATDPTIVGVNDKTDGSONI1SNASCTT NCLAPLAKVLHDPGIVKGLMTHAYTQDQLODGPESLDRARAAALNVVPTSGA AKAGLVMPKSLKCLQYALRVPIPTGSLVTLTADLSKCVSNELNAPQDAAGELX GILKYVDAPKESDSDIVTDPHSS1PDSGLTKVTAGQAKVSVWIDNENYSNRLVVLGK VGSKL"
6892..7902
/gene="gap"
/note="Pfam match to entry PF00044 gpdh, glyceraldehyde 3-phosphate dehydrogenases, score 733.80, E-value 7.9e-219"
7354..7377
misc_feature

```

```

/gene="gap"
/note="PF00071 Glyceraldehyde 3-phosphate dehydrogenase active site"
7911..9161
/gene="pgk"
/note="synonym: ML0571"
7911..9161
/gene="pgk"
/EC_number="2.7.2.3"
/note="Similar to M. tuberculosis pgk, Rv1437, phosphoglycerate kinase, SW:PGK_MYCTU (006821) (412 aa); Pista score E(): 0, 90.6% identity in 412 aa overlap. Similar to many e.g. Corynebacterium glutamicum pgk, phosphoglycerate kinase, SW:PGK_CORGL (Q01655) (403 aa); Pista score E(): 0, 59.2% identity in 407 aa overlap. Previously sequenced as SW:PGK_MYCLE (P46712) (416 aa); Pista score E(): 0, 99.8% identity in 416 aa overlap. Contains pfam match to entry PF00162 pgk, phosphoglycerate kinases. Contains PF00111 Phosphoglycerate kinase signature."
/codon_start=1
/transl_table=11
/product="phosphoglycerate kinase"
/protein_id="CAC30079.1"
/db_xref="GI:13092769"
/db_xref="GOA:P46712"
/db_xref="SWISS-PROT:P46712"
/translation="MKIPNLKDLLEGVGRVLRCTLNVPDGDGAIITDLGRVTAS VPTKALLRAGAKIVVAHILGRPKNGDPKLSLEPVAALGEQLQGNVQVLCSTDRSP VQDNACVLERITDGLLILLONIRPDRETSTKDDERLAKQLVHVBAGAPVSDG PCVHREQASVDVATLPHYAGILVADEIRILLBOLTSSAKRPVAVLGGSKSVKDLG VITSLKADSVLVLGGCGFTPLAQGVCKSLKLETEMETCESLDITADVLRLLPM DVIYKRFVADSPQTVAADRIPOGLKSLDGPESYKRPATLSNASTIPNPGMPGVF EPLAAGATRGVAVALIVATGKAFSVVGGGSAATKAUSLPFGSVSHLSTGGASL EYLEGKTLPGIEVLGRBQPRGDS"
7911..9140
/gene="pgk"
/note="Pfam match to entry PF00162 pgk, phosphoglycerate kinases, score 757.00, E-value 7.9e-224"
7962..7994
/gene="pgk"
/note="PF00111 Phosphoglycerate kinase signature"
9149..9153
/gene="pgk"
/note="possible RBS"
9158..9943
/gene="tpi"
/note="synonym: ML0572"
9158..9943
/gene="tpi"
/EC_number="5.3.1.1"
/note="Similar to M. tuberculosis tpi, Rv1438, triosephosphate isomerase, SW:TPIS_MYCTU (008408) (261 aa); Pista score E(): 0, 83.9% identity in 261 aa overlap. Similar to many e.g. Corynebacterium glutamicum tpi, triosephosphate isomerase, SW:TPIS_CORGL (P19583) (259 aa); Pista score E(): 0, 64.5% identity in 256 aa overlap. Previously sequenced as SW:TPIS_MYCLE (P46711) (261 aa); Pista score E(): 0, 99.6% identity in 261 aa overlap. Contains pfam match to entry PF00121 TIM, triosephosphate isomerase. Contains PF00171 Triosephosphate isomerase active site."
/codon_start=1
/transl_table=11
/product="triosephosphate isomerase"
/protein_id="CAC30083.1"
/db_xref="GI:13092770"
/db_xref="GOA:P46711"
/db_xref="SWISS-PROT:P46711"
/translation="MSRKSIIAGNPMONTNHPFAIALVQKIAFSLPKYKDYDVTVL PPFTDLSQVTLVDGDKRLTYGAQLDSGAYTGDIISGAFKAGSVLVGHSZ RETHYDGDALVAAKTAAALKNSLTPIVCIGBYLEIRVGEVSHCNQLRGSLAGLS PEQIGNVIVYEPVWAIQTURVAAADAQSVCEAIRKELGALASQVAZVRVLYGGS"

```


gene
CDS

```

Query Match      61.8%; Score 20.4; DB 1; Length 345783;
Best Local Similarity 80.0%; Pred. No. 9.1e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 CATCTCTCTCCGAGCCGCTCGAAATAGTG 30
18290 CATCTGTTGATGCGCGCTGCATATAGTG 18319

RESULT 26
AC100516          61580 bp      DNA      linear      HTG 22-NOV-2001
AC100516          Mus musculus clone RP23-145H2, LOW-PASS SEQUENCE SAMPLING.
AC100516          Mus musculus
AC100516.1        GI:17047882
HTG1, HTG5 PHASE0
Mus musculus (house mouse)
Mus musculus
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 61580)
Barnes, B., Kinton, L., Nusbaum, C. and Lander, E.

```

Unpublished

2 (bases 1 to 61590)

JOURNAL
REFERENCE
AUTHORS

Birren, B., Linton, L., Bastien, V., Boguslavsky, L., Bouckgaert, R., Anderson, S., Barnes, N., Campolano, A., Chang, J., Chazaro, B., Brown, A., Camarata, G., Collins, S., Collamore, A., Cook, A., Choe, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Dias, J. S., Dodge, S., Faro, S., Ferrel, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, N., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kellis, C., Lakocque, K., Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Maneus, L., Mihova, T., Mclennan, J., Murphy, T., Naylor, J., Nguyen, C., Norum, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Ribick, N., Riley, R., Rise, C., Rogov, P., Roman, J., Ronetti, M., Roy, A., Santos, R., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J., Topham K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zambek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (22-Nov-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

JOURNAL

All repeats were identified using RepeatMasker:

COMMENT

Snit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIZB
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L5516
 Center clone name: i45_M_2

* NOTE: This record contains 76 individual sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 703: contig of 703 bp in length
* 704 803: gap of 100 bp
* 804 1515: contig of 712 bp in length
* 1516 1615: gap of 100 bp
* 1616 2199: contig of 584 bp in length
* 2200 2299: gap of 100 bp
* 2300 2398: contig of 699 bp in length
* 2399 3098: gap of 100 bp
* 3099 3797: contig of 699 bp in length
* 3798 3897: gap of 100 bp
* 3898 4615: contig of 718 bp in length
* 4616 4716: gap of 100 bp
* 4716 5421: contig of 706 bp in length
* 5422 5521: gap of 100 bp
* 5522 6223: contig of 702 bp in length
* 6224 6323: gap of 100 bp
* 6324 7028: contig of 705 bp in length
* 7029 7128: gap of 100 bp
* 7129 7830: contig of 702 bp in length
* 7831 7930: gap of 100 bp
* 7931 8644: contig of 714 bp in length
* 8645 8744: gap of 100 bp
* 8745 9427: contig of 683 bp in length
* 9428 9527: gap of 100 bp
* 9528 10265: contig of 738 bp in length
* 10266 10365: gap of 100 bp
* 10366 11100: contig of 735 bp in length
* 11101 11200: gap of 100 bp
* 11201 11938: contig of 738 bp in length
* 11939 12038: gap of 100 bp
* 12039 12740: contig of 702 bp in length
* 12741 12840: gap of 100 bp
* 12841 13544: contig of 704 bp in length
* 13545 13644: gap of 100 bp
* 13645 14311: contig of 667 bp in length
* 14312 14411: gap of 100 bp
* 14412 15133: contig of 722 bp in length
* 15134 15233: gap of 100 bp
* 15234 15958: contig of 725 bp in length
* 15959 16058: gap of 100 bp
* 16059 16747: contig of 689 bp in length
* 16748 16847: gap of 100 bp
* 16848 17583: contig of 736 bp in length
* 17584 17683: gap of 100 bp
* 17684 18404: contig of 721 bp in length
* 18405 18504: gap of 100 bp
* 18505 19228: contig of 724 bp in length
* 19229 19328: gap of 100 bp
* 19329 20004: contig of 676 bp in length
* 20005 20104: gap of 100 bp
* 20105 20835: contig of 711 bp in length
* 20836 20935: gap of 100 bp
* 20936 21610: contig of 695 bp in length
* 21611 21710: gap of 100 bp
* 21711 22444: contig of 734 bp in length
* 22445 22544: gap of 100 bp
* 22545 23250: contig of 706 bp in length
* 23251 23350: gap of 100 bp
* 23351 24065: contig of 715 bp in length
* 24066 24165: gap of 100 bp
* 24166 24916: contig of 751 bp in length
* 24917 25016: gap of 100 bp
* 25017 25737: contig of 721 bp in length
* 25738 25837: gap of 100 bp

25838 26541: contig of 704 bp in length
* 26542 27311: contig of 690 bp in length
* 27312 27411: gap of 100 bp
* 27412 28129: contig of 698 bp in length
* 28130 28229: gap of 100 bp
* 28230 28982: contig of 753 bp in length
* 28983 29082: gap of 100 bp
* 29083 29792: contig of 710 bp in length
* 29793 29892: gap of 100 bp
* 29893 30614: contig of 722 bp in length
* 30615 30714: gap of 100 bp
* 30715 31454: contig of 740 bp in length
* 31455 31554: gap of 100 bp
* 31555 32255: contig of 711 bp in length
* 32256 32355: gap of 100 bp
* 32356 33058: contig of 703 bp in length
* 33059 33158: gap of 100 bp
* 33159 33875: contig of 707 bp in length
* 33876 33975: gap of 100 bp
* 33976 34691: contig of 716 bp in length
* 34692 34791: gap of 100 bp
* 34792 35515: contig of 724 bp in length
* 35516 35615: gap of 100 bp
* 35616 36311: contig of 696 bp in length
* 36312 36411: gap of 100 bp
* 36412 37119: contig of 708 bp in length
* 37120 37219: gap of 100 bp
* 37220 37889: contig of 670 bp in length
* 37890 37989: gap of 100 bp
* 37990 38725: contig of 736 bp in length
* 38726 38825: gap of 100 bp
* 38826 39512: contig of 687 bp in length
* 39513 39612: gap of 100 bp
* 39613 40340: contig of 728 bp in length
* 40341 40440: gap of 100 bp
* 40441 41167: contig of 727 bp in length
* 41168 41267: gap of 100 bp
* 41269 41956: contig of 729 bp in length
* 41957 42056: gap of 100 bp
* 42057 42807: contig of 711 bp in length
* 42808 42907: gap of 100 bp
* 42908 43628: contig of 721 bp in length
* 43629 43728: gap of 100 bp
* 43729 44438: contig of 710 bp in length
* 44439 44538: gap of 100 bp
* 44539 45280: contig of 742 bp in length
* 45281 45380: gap of 100 bp
* 45381 46132: contig of 752 bp in length
* 46133 46232: gap of 100 bp
* 46233 46936: contig of 704 bp in length
* 46937 47036: gap of 100 bp
* 47037 47750: contig of 714 bp in length
* 47751 48574: gap of 100 bp
* 48575 48674: gap of 100 bp
* 48675 49408: contig of 734 bp in length
* 49409 49508: gap of 100 bp
* 49509 50220: contig of 722 bp in length
* 50221 50320: gap of 100 bp
* 50321 51038: contig of 718 bp in length
* 51039 51138: gap of 100 bp
* 51139 51856: contig of 716 bp in length
* 51857 52658: gap of 100 bp
* 52659 52758: contig of 702 bp in length
* 52759 53459: contig of 711 bp in length
* 53470 53569: gap of 100 bp
* 53570 54279: contig of 710 bp in length
* 54280 54379: gap of 100 bp
* 54380 55108: contig of 729 bp in length
* 55109 55208: gap of 100 bp
* 55209 55917: contig of 709 bp in length

Mapping information for this clone was provided by Dr. John D. McClellan, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Maratton, N., Canas, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Genoscreen Research Genetics, Inc. (<http://www.researchgen.com>) or Pter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of RP11-123G24;
actual end is at base position 161955 of RP11-123G24.

Region	Location/Qualifiers
region	1. .161955
region	/organism="Homo sapiens"
region	/mol_type="genomic DNA"
region	/db_xref="taxon:9606"
region	/chromosome="2"
region	/map="2"
region	/clone="RP11-121G24"
region	/clone_lib="RPC1-11"
region	1. .72
region	/rpt_family="Alu"
region	71. .119
region	/rpt_family="Mariner"
region	318. .437
region	/rpt_family="Alu"
region	447. .478
region	/rpt_family="(TAAA)n"
region	603. .662
region	/rpt_family="MIR"
region	1018. .1077
region	/rpt_family="AT-rich"
region	1183. .1489
region	/rpt_family="Alu"
region	2978. .3023
region	/rpt_family="CT-rich"
region	3476. .3517
region	/rpt_family="(TTTTG)n"
region	3492. .3796
region	/rpt_family="Alu"
region	4123. .4189
region	/rpt_family="AT-rich"
region	4389. .4411
region	/rpt_family="AT-rich"
region	4975. .5354
region	/rpt_family="MaLR"
region	6933. .7303
region	/rpt_family="MaLR"
region	8280. .8313
region	/rpt_family="AT-rich"
region	8877. .8897
region	/rpt_family="AT-rich"
region	9568. .9982
region	/rpt_family="L1"
region	9987. .10470
region	/rpt_family="L1"
region	10166. .10196
region	/rpt_family="AT-rich"
region	10471. .10772
region	/rpt_family="Alu"
region	10773. .10818
region	/rpt_family="L1"
region	11492. .11794
region	/rpt_family="Alu"
region	11768. .11815
region	/rpt_family="(TAAA)n"

```

* 55918 56017: gap of 100 bp

Query Match
Best Local Similarity 61.2%; Score 20.3; DB 2; Length 61580;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 CATCTCTTCTCCGAGCCGGTCGGAANTAGTGAGT 33
|||||
5985 CTTCTCTTCTGAGGTCATCCAACTACTTAGT 6017

SULT 27
068718/c AC068718 161955 bp DNA linear PRI 09-JAN-2002
FFINITION Homo sapiens BAC clone RP11-123G24 from 2, complete sequence.

REGION	AC068718.6	GI:14718377
WORDS	HTG.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

REFERENCE
AUTHORS
TITLE
JOURNAL
GENOME RES. 8 {11}, 1097-1108 (1998)
MEDLINE
9847074
PUBMED

3 (bases 1 to 161955)
 Unpublished (2001)
 The sequence of Homo sapiens BAC clones RP11-123C24
 Du, Y., Meyer, R., Doeberer, A. and Dignan, G.
 2000, 1 to 161955

TITLE Direct Submission
JOURNAL Submitted (07-MAY-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA

Waterston, R.H.
Direct Submission
 Submitted (14-JUL-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 161955)

TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Jul 14, 2001 this sequence version replaced gi114573737.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics -----
Center project name: H MH0123G24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

```

repeat_region 11880..11984
/rpt_family="L1"
repeat_region 12013..12038
/rpt_family="(T)n"
repeat_region 12022..12345
/rpt_family="SVA"
repeat_region 13875..13926
/rpt_family="(CAAA)n"
repeat_region 13927..13947
/rpt_family="(GGGAGA)n"
repeat_region 13966..14618
/rpt_family="L1"
misc_feature 14609..14953
/notes="similar to Homo sapiens EST AW977707
(NID:g8169086)"
misc_feature 14610..14862
/notes="similar to Homo sapiens EST AA193339 (NID:gl782929)
x41b10.r1"
misc_feature 14613..14844
/notes="similar to EST BU185210 (NID:gl4659619)"
repeat_region 14952..15347
/rpt_family="L1"
repeat_region 16972..17368
/rpt_family="MALR"
repeat_region 17449..17471
/rpt_family="AT_rich"
repeat_region 17648..17852
/rpt_family="MER1_type"
repeat_region 18846..18960
/rpt_family="MIR"
repeat_region 19246..19371
/rpt_family="L2"
repeat_region 19899..19806
/rpt_family="MIR"
repeat_region 21648..21669
/rpt_family="AT_rich"
repeat_region 21689..21919
/rpt_family="L1"
repeat_region 22846..22893
/rpt_family="(CA)n"
repeat_region 22900..22938
/rpt_family="(CAT)n"
repeat_region 25287..25335
/rpt_family="AT_rich"
repeat_region 25473..25502
/rpt_family="AT_rich"
repeat_region 25932..25952
/rpt_family="AT_rich"
repeat_region 26136..26231
/rpt_family="AT_rich"
repeat_region 26679..26810
/rpt_family="L2"
repeat_region 27642..27671
/rpt_family="AT_rich"
repeat_region 27661..27822
/rpt_family="Alu"
repeat_region 28513..28723
/rpt_family="ERV1"
repeat_region 29254..29386
/rpt_family="ACHobo"
repeat_region 29851..29905
/rpt_family="AT_rich"
repeat_region 30015..30084
/rpt_family="MIR"
repeat_region 30352..30677
/rpt_family="ERV1"
repeat_region 30678..31255

```

Query Match 61.2%; Score 20.2; DB 9; Length 161955;
Best Local Similarity 75.8%; Pred. No. 1a+03;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

DB	74467	CATCTCTTCACGGCCGAGGAGGACAGGTGAT	74335
RESULT 28			
AL626782			
LOCUS			
DEFINITION			
AL626782			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

AL626782 177264 bp DNA linear ROD 29-JUN-2002
 Mouse DNA sequence from clone RP23-335N15 on chromosome X, complete
 sequence.
 AL626782
 AL626782.10 GI:20068573
 HTG
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 177264)
 Phillimore, B.
 Direct Submission
 Submitted (29-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 hammy@esanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:19031574.
 During sequence assembly data is compared from overlapping clones.
 Where differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all
 regions were either double-stranded or sequenced with an alternate
 chemistry or covered by high quality data (i.e., phred quality >=
 30); an attempt was made to resolve all sequencing problems, such
 as compressions and repeats; all regions were covered by at least
 one plasmid subclone or more than one M13 subclone; and the
 assembly was confirmed by restriction digest. The following
 abbreviations are used to associate primary accession numbers given
 in the feature table with their source databases: Em: EMBL; Sw:
 SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information on the WormPep
 database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-335N15 is
 from the RPI-33 Mouse PAC library
 constructed by the group of Peter de Jong.
 For further details see <http://www.chori.org/bacpac/home.htm>.
 VECTOR: pBACe3.6
 ----- Genome Center
 Center: UK Medical Research Council
 Contact code: UK-MRC
 Web site: <http://mrcseq.har.mrc.ac.uk>
 Contact: mouse@har.mrc.ac.uk

```

-----
FEATURES
  source          location/Qualifiers
    1..177264
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /chromosome="X"
      /clone="RP23-335N15"
      /clone_lib="RPCI-23"

ORIGIN
Query Match          61.2%;  Score 20.2;  DB 10;  Length 177264;
Best Local Similarity 75.8%;  Pred. No. 1,1e+03;
Matches 25;  Conservative 0;  Mismatches 8;  Indels 0;  Gaps 0;

Qy      1  CATCTTCTTCGAGCGCGGTGGAATAGTGACT 33
          ||| ||| ||| ||| ||| ||| ||| |||
Db 144190 CTTCTCTCTCTGAGCTCATCCCAATACTTACT 144222
          ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 29			
AC110030			
LOCUS	AC110030	207051 bp	DNA linear
			HTG 06-MAY-2002

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus clone RP23-49B18, WORKING DRAFT SEQUENCE, 72 ordered
pieces.
AC110030
GI:20455635
HTG, HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
Mus musculus (house mouse)
Mus musculus
Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 207051)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bouguelvsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.-S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Illiev, I., Johnson, K., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, K.,
Lander, E., Lechoczy, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McSwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Ming, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 207051)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bouguelvsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.-S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Illiev, I., Johnson, K., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, K.,
Lander, E., Lechoczy, J., Levine, R., Liu, G., MacLean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McSwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Ming, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (09-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 207051)

Birken, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Bouguelvsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.-S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Illiev, I., Johnson, K., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Lander, E., Lechoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McSwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Ming, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 6, 2002 this sequence version replaced gi:20389485.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIR
Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22336
Center clone name: 49_B_18

----- Summary Statistics
Sequencing Vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus Quality: 186564 bases at least Q40
Consensus Quality: 195257 bases at least Q30
Consensus Quality: 198001 bases at least Q20
Insert size: 180000; agarose-fp
Insert size: 199931; sum-of-contigs
Quality coverage: 6.1 in Q20 bases; agarose-fp
Quality coverage: 7.3 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 72 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces
is believed to be correct as given, however the sizes
of the gaps between them are based on estimates that have
been provided by the submitter.
This sequence will be replaced
by the finished sequence as soon as it is available and
the accession number will be preserved.
5591: contig of 5591 bp in length
5592: gap of 100 bp
5593: contig of 889 bp in length
5594: gap of 100 bp
5595: contig of 785 bp in length
5596: gap of 100 bp
5597: contig of 1028 bp in length
5598: gap of 100 bp
5599: contig of 638 bp in length
5600: gap of 100 bp
5601: contig of 922 bp in length
5602: gap of 100 bp
5603: contig of 608 bp in length
5604: gap of 100 bp
5605: contig of 989 bp in length
5606: gap of 100 bp
5607: contig of 666 bp in length
5608: gap of 100 bp
5609: contig of 717 bp in length
5610: gap of 100 bp
5611: contig of 948 bp in length
5612: gap of 100 bp
5613: contig of 918 bp in length
5614: gap of 100 bp
5615: contig of 814 bp in length
5616: gap of 100 bp
5617: contig of 1079 bp in length
5618: gap of 100 bp
5619: contig of 715 bp in length
5620: gap of 100 bp
5621: contig of 644 bp in length
5622: gap of 100 bp
5623: contig of 669 bp in length
5624: gap of 100 bp
5625: contig of 948 bp in length
5626: gap of 100 bp
5627: contig of 931 bp in length
5628: gap of 100 bp
5629: contig of 713 bp in length
5630: gap of 100 bp
5631: contig of 709 bp in length
5632: gap of 100 bp
5633: contig of 724 bp in length
5634: gap of 100 bp
5635: contig of 561 bp in length
5636: gap of 100 bp
5637: contig of 784 bp in length
5638: gap of 100 bp

```

* 26393 27188: contig of 796 bp in length
* 27189 gap of 100 bp
* 27389 27935: contig of 647 bp in length
* 27936 28035: gap of 100 bp
* 28036 28726: contig of 691 bp in length
* 28727 28826: gap of 100 bp
* 28827 29705: contig of 879 bp in length
* 29706 29805: gap of 100 bp
* 29806 30503: contig of 698 bp in length
* 30504 31205: gap of 100 bp
* 31206 31305: contig of 602 bp in length
* 31306 32280: contig of 975 bp in length
* 32281 32380: gap of 100 bp
* 32381 33338: contig of 958 bp in length
* 33339 34332: gap of 100 bp
* 34333 34732: contig of 1194 bp in length
* 34733 35215: contig of 1183 bp in length
* 35216 36016: gap of 100 bp
* 36016 37046: contig of 1031 bp in length
* 37047 37146: gap of 100 bp
* 37147 37820: contig of 674 bp in length
* 37821 38598: gap of 100 bp
* 38599 39348: contig of 678 bp in length
* 39349 40370: contig of 650 bp in length
* 40371 40779: gap of 100 bp
* 40779 41750: contig of 922 bp in length
* 41751 41850: gap of 100 bp
* 41851 42574: contig of 1280 bp in length
* 42575 44058: gap of 100 bp
* 44059 45232: contig of 724 bp in length
* 45233 45332: contig of 100 bp
* 45333 45980: contig of 1383 bp in length
* 45981 46000: gap of 100 bp
* 46001 50266: contig of 1075 bp in length
* 50267 50627: gap of 100 bp
* 50628 52357: contig of 1647 bp in length
* 52358 53675: contig of 1421 bp in length
* 53676 53775: gap of 100 bp
* 53776 55165: contig of 1926 bp in length
* 55166 56961: contig of 1631 bp in length
* 56962 57061: gap of 100 bp
* 57062 58728: contig of 1696 bp in length
* 58729 59622: gap of 100 bp
* 59623 59762: contig of 1667 bp in length

```

```

Query Match 61.2%; Score 20.2; DB 2; Length 207051;
Best Local Similarity 75.8%; Pred. No. 1.1e+03;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1 CATCTCTTCGAGCCGCTCGAATAGTGAGT 33
|||||
DB 133148 CATCTCTTCGAGCCGCTCGAATAGTGAGT 133180
|||||

```

```

RESULT 30
AC122907 AC122907 233904 bp DNA linear ROD 10-DEC-2003
LOCUS Mus musculus chromosome 15 clone RP23-4914, complete sequence.
DEFINITION AC122907
ACCESSION AC122907
VERSION AC122907.4 GI:39652681
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 233904)
McPherson, J.D. and Waterston, R.H.
DIRECT SUBMISSION
JOURNAL Submitted (28-MAY-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
AUTHORS 3 (bases 1 to 233904)
Wilson, R.K.
DIRECT SUBMISSION
JOURNAL Submitted (30-OCT-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
AUTHORS 4 (bases 1 to 233904)
Wilson, R.K.
DIRECT SUBMISSION
JOURNAL Submitted (10-DEC-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Dec 10, 2003 this sequence version replaced gi:38044277.

```

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
----- Project Information -----
Center project name: M_BA0049104

```

FEATURES

```

source
1. .233904
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="15"
/clone="RP23-4914"

```

ORIGIN

```

Query Match 61.2%; Score 20.2; DB 10; Length 233904;
Best Local Similarity 75.8%; Pred. No. 1.1e+03;
Matches 25; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 CATCTCTTCGAGCCGCTCGAATAGTGAGT 33
|||||
DB 161015 CATCTCTTCGAGCCGCTCGAATAGTGAGT 161047
|||||

```

```

RESULT 31
AP003581 348050 bp DNA linear BCT 28-NOV-2001
LOCUS Nostoc sp. PCC 7120 DNA, complete genome, section 1/19.
DEFINITION AP003581 BA000019
ACCESSION AP003581.1 GI:117134979
VERSION
KEYWORDS
SOURCE Nostoc sp. PCC 7120
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
ORGANISM

```

```

REFERENCE
AUTHORS Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S.,
Maranabe, A., Iriuch, M., Ishikawa, A., Kawashina, K., Kimura, T.,
Kishida, Y., Kohara, M., Matsumoto, M., Matsumoto, A., Muraki, A.,
Nakasaki, N., Shimpo, S., Sugimoto, M., Takazawa, K., Yamada, M.,
Yasuda, M. and Tabata, S.
Complete genomic sequence of the filamentous nitrogen-fixing
Cyanobacterium Anabaena sp. strain PCC 7120
DNA Res. 8 (5), 205-213 (2001)
JOURNAL 21595285
MEDLINE 11759840
PUBMED
REFERENCE 2 (bases 1 to 348050)
Kaneko, T.
DIRECT SUBMISSION
JOURNAL Submitted (02-MAY-2001) Takakazu Kaneko, Karusa DNA Research

```

Institute, The First Laboratory for Plant Gene Research, Yana
1532-3, Kisarazu, Chiba 292-0812, Japan

(E-mail: kaneko@kcasu.az.jp,
URL: http://www.kasusa.or.jp/cyanebase/

Tel: 81-438-52-3935 (ex. 2338), Fax: 81-438-52-3934

FEATURES

```

source
    1. 348050
    /organism="Nostoc sp. PCC 7120"
    /mol_type="genomic DNA"
    /db_xref="taxon:103690"
    /note="synonym: Anabaena sp. PCC7120"
    /complement(981..1718)
    /gene="all0002"
    /complement(981..1718)
    /gene="all0002"
    /note="ORF ID: all0002"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB77526.1"
    /db_xref="GI:17134980"
    /translation="MLFTNQSIENKGIKMMKPIIFLTQLLTSLALFVSVNVA
    QDTPASFPQMAKICTSDVENLIPFLQGRGNDLSYLAQSGFTSDHGSVCY
    VNSKKEGYTLTKVQVNGKLGISPLDGTITABQDSRLDFFNTLIANTHIEA
    DNRQSRVYLEAPILVAGKIOPSRGVFPQPNRALVIYHKLITGELQGTALINI
    QVPKSVSSGSLNKLIPVSPQGVKPNQ"
    /complement(2617..2805)
    /gene="as10003"
    /complement(2617..2805)
    /note="ORF ID: as10003"
    /codon_start=1
    /transl_table=11
    /protein_id="BAB77527.1"
    /db_xref="GI:17134981"
    /translation="MDTPLEKQWTLTPVAALIPMFNVGVVLRNDRMLGSRNPT
    RVLTSTIVQQLVLP"
    /complement(2858..3247)
    /gene="ssrA"
    /complement(2858..3247)
    /note="ssrA"
    /product="10Ss RNA"
    /complement(3418..4365)
    /gene="atpC"
    /complement(3418..4365)
    /note="ORF ID: all0004"
    /codon_start=1
    /transl_table=11
    /product="ATP synthase subunit gamma"
    /protein_id="BAB77528.1"
    /db_xref="GI:17134982"
    /translation="WPNLESIRDRIQSVNKKITAMELVAAARVERAQEVIATRP
    PADRLAQLVGLQTLRFEDVDLFLKQEVKSVGLVIGDGLCGYNTYIRRAE
    NAKELKAEGLDYTVIVGRKAEQVFRRRQPIDASVTLGKQIFPADANKIADLLS
    LFLSEKVDRELVTVPVSVSVSRPVITCLPLDTQGLEADDSEIPLTTRGQPOVE
    RQVTSQARLPEDMLPEQDPVQLDLSLLPLYSNQLRALQESAELAAARTAMSN
    ASERAGSLRGLSLSYNKAQKATIQHLLVVGNEALT"
    /complement(4454..5974)
    /gene="atpA"
    /complement(4454..5974)
    /note="atpA"
    /note="ORF ID: all0005"
    /codon_start=1
    /transl_table=11
    /product="ATP synthase subunit alpha"
    /protein_id="BAB77529.1"
    /db_xref="GI:17134983"
    /translation="NSISIRPDEISSIIQQQIEQVQVQVKNVGVTVLQVGDGIARIY
    GLEKAMAGELLSEFDGTGVIQNLNEDNVGAVLMGEGREIQEGSTVTATGRIAQIGV
  
```

```

    /complement(6138..6689)
    /gene="atpD"
    /complement(6138..6689)
    /note="ORF ID: all0006"
    /codon_start=1
    /transl_table=11
    /product="ATP synthase subunit delta"
    /protein_id="BAB77530.1"
    /db_xref="GI:17134984"
    /translation="MTSKVANTEVAQPYAQLLSIAKSKSLTEFEGTDARTLLNLTLE
    NQQLANFDIDNPPFAENKKAIIKQLSBSAPVLRNELLVLDKERIFFELPILQVLA
    LRQLNGTVLAEVTSVAULTEDQQQAVTEKVLATKAGQVELATKYVDSLDGGVIKY
    GSQVDSISIKQQLRSLSLSNS"
    /complement(6686..7249)
    /gene="atpD"
    /complement(6686..7249)
    /note="ORF ID: all0007"
    /codon_start=1
    /transl_table=11
    /product="ATP synthase subunit b"
    /protein_id="BAB77531.1"
    /db_xref="GI:17134985"
    /translation="WGTFLLEMAASAVGZELAEAGEGGFCGLNTNLTNLTNINLII
    ITVLFPVKRVLGNTLTRENTETAIKNAQRQAADAKQLKBAQKLECAQAEARI
    KSKAQDSNADQAGQAIQAQAVDIERLQBAADLNAELDEAIAQRQVRVALQKVE
    SELQGISDQAKTLIDRSIAQLGGV"
    /complement(7359..7850)
    /gene="atpG"
    /complement(7359..7850)
    /note="atpG"
    /note="ORF ID: all0008"
    /codon_start=1
    /transl_table=11
    /product="ATP synthase subunit b"
    /protein_id="BAB77532.1"
    /db_xref="GI:17134986"
    /translation="MTHTWTLTAAVEKVAKEGGLPDLDTATPLMATQFLLALILNATL
    YKPLGKIDGRNVRNNQLEAQRSLSKAKLAEAVEQELAGARQACTITADAQABA
    OKIAAEKVAAKQAQAEQAEQAEIQKQKQALASLEQQVDALSRQLKLGADLV
    KQR"
    /complement(7996..8241)
    /gene="atpH"
    /complement(7996..8241)
    /note="atpH"
    /note="ORF ID: as1009"
    /codon_start=1
    /transl_table=11
    /product="ATP synthase subunit c"
    /protein_id="BAB77533.1"
    /db_xref="GI:17134987"
    /translation="NDPINSAAVLAALAVGLAIGPGICGGNAGQAVEGIAQPE
    AEGKIRGTLISLAPNEALITYGLVVALVLLPANPFA"
    /complement(8324..9079)
    /gene="atpI"
    /complement(8324..9079)
    /note="atpI"
    /codon_start=1
    /transl_table=11
  
```



```

/functions="modulation protein"
/functions="transcriptional regulator"
/notes="modulation protein; transcriptional regulator; a
longer form than Rhizobium fredii NodB2 (encoded by
Genbank Accession Number L38460) may exist; possible
alternative start point at nucleotide 492755"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="NodB2"
/protein_id="AAB91901.1"
/db_xref="GI:2182715"

Query Match      60.1%; Score 20; DB 1; Length 10140;
Best Local Similarity 82.1%; Pred. No. 1e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CTCCTCCGACCGCGTGCAGTAATAGTCA 31
Db 4483 CTCCTCCGACCGCGTGCAGTCTCTCA 4510

RESULT 35
AE000102 Rhizobium sp. NGR234 10242 bp DNA linear BCT 12-DEC-1997
LOCUS Rhizobium sp. NGR234 plasmid pNGR234a, section 39 of 46 of the
DEFINITION complete plasmid sequence.
ACCESSION AE000102 U00090
KEYWORDS
SOURCE Rhizobium sp. NGR234
ORGANISM Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
REFERENCE 1 (bases 1 to 10242)
AUTHORS Freiberg, C., Fellay, R., Broughton, M.J., Rosenthal, A.
and Perret, X.
TITLE Molecular basis of symbiosis between Rhizobium and legumes
JOURNAL Nature 387 (6631), 394-401 (1997)
MEDLINE 97305956
PubMed 9163424
REFERENCE 2 (bases 1 to 10242)
AUTHORS Freiberg, C.
TITLE Direct Submission
JOURNAL Submitted (22-NOV-1996) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 10242)
AUTHORS Freiberg, C.
TITLE Direct Submission
JOURNAL Submitted (12-DEC-1997) Genome Analysis, Institute for Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REMARK Update by submitter
FEATURES
    location/Qualifiers
    1..10242
        /organism="Rhizobium sp. NGR234"
        /mol_type="genomic DNA"
        /strain="NGR234"
        /db_xref="taxon:394"
        /clone="pXB368"
        /plasmid="pNGR234a"
        /size="1600"
        /genes="nifD1"
        /size="1600"
        /genes="nifD1"
        /size="1600"
        /genes="nifD1"
        /note="nitrogenase Mo-Fe protein alpha chain (nitrogenase
        component 1)"
        /codon_start=1
        /transl_table=11
        /product="NifD1"
        /protein_id="AAB91901.1"
        /db_xref="GI:2182681"
        /translation="MSLDYDNDLHQLITVLSQYHKAARQKHLSVASDEAV
        GEGHTLSECDVKNIXISFGVMTIRGCYAGSGVGVGVDVHLSHGPGVGCQYIS
        WGRNRYTVGTGVTFTVTOFTSDFOBKDFYFGDKQCLEQVIDRIEELPPLNGCTTI
        QSECTIGLIGDDIIVASRKAARHETTVVRCGEGFRGVSQSLGHIAIDATRDVDFD
        KADGTQVFTGTVNVDYGVNIGDANASRIILREIGLAVVGNWSDATLAEVFR
        APTAKLNIHCVRYNNTYICRMERYVAPMKNYVFPQSOIILASLRKARHFGPTIBE
        RARVATYRPLVDAVDKVPRLQGVNLYVGLRPRHVTAVEDLGMQVGTGYE
        PARNDYORTGHYVTKTLLYDDATSYELDTPIERLPDLYSGSKERYPVQGMGIPP
        RQHSWQYSGPVHGVGDFALFARDMDLAINNPVMDLYDAPKKNTVFAVAAL"
        /size="1696"
        /genes="nifK1"
        /size="1696"
        /genes="nifK1"
        /note="nitrogenase Mo-Fe protein beta chain (nitrogenase
        component 1)"
        /codon_start=1
        /transl_table=11
        /product="NifK1"
        /protein_id="AAB91901.1"
        /db_xref="GI:2182682"
        /translation="MPASAKILDEAPLPRPEYROMLAERKLNSECHEPRLVTDOR
        EYKGNWETERNLALVNPAPAKCOPLGAVPAAGPRTMSFVHGSGCCVATYRSH
        LSRHTEPASAVSSMTEDAVYVGLNNDVGLANTYALDPMIAVSTI QAVVIGD
        DLHGPIENAKGEGAVPEPVPVPAHTPAPVGHVDTGDSKVGILHHPWQQTQAA
        GTINIPGDFGCVGNRELRLTLGVSYTFIQDASQDFPDPDGRVMTDGGTTI
        KALRAALAEATLSLHYNRSKRLTETCRVQATAPAPHYPLGINATDPLKVSYSIG
        RIPTTILERGLVDAMDSQSLGKTYAYIGDPOFVYVAMAFVMEGEPKCLER
        TNGTAAQAMTELLASPPCKQAKVWPKOLWALSLFTPEVDLLIGNSYKCTLER
        DTGPTLRMPPIFDHRRHPPPLMCYQCCLELLTITDITFDLDRQETMQATVDSY
        YDLYT"
        /size="3293"
        /genes="nifH"
        /size="3293"
        /genes="nifH"
        /note="nitrogenase Fe-Mo cofactor biosynthesis protein
        NifH"
        /codon_start=1
        /transl_table=11
        /product="NifH"
        /protein_id="AAB91902.1"
        /db_xref="GI:2182683"
        /translation="MSSLVKVIQDVDEPACERNRKNDSNARKSGCSFELTPGAANG
        CAPDGAIVLPITDVAHLVHAPLACGNSWDRGTASSGPTLNRTSFTTDLTEQDII
        MGNSEKLLPALKIQRREAYQPPALFYATATGIDGDEIAICKRAEETCGLPVPIIN
        SPQVGSKNLKGKAGEALDORVIGTVPDHAGCPDINILGEPNLSGFWPLPFLER
        LGIRVACTPGDARYLDIASARARAAVVCSTALINLARKMEERMDIPFFGSGYGI
        TATSALQIADLLVKGDTLEITDRTDALIAEERAAWKLEFYRPLKGRVLIPT
        GQVGSLSVHALMIGHEIVGTSVKSTVEDKERIQVLCODLQMFSPMSPRELYAML
        SERKADIMLSGRTQPTALKAKMEVDINQRHRPVAGYDGTVELVQIDLAIRNPVW
        SOVREPAWHSRPTQRPPLNTRPDTATSGVRSIRRHADAGDFVEC"
        /size="4793"
        /genes="nifN"
        /size="4793"
        /genes="nifN"
        /note="nitrogenase Fe-Mo cofactor biosynthesis protein
        NifN"
        /codon_start=1
        /transl_table=11
        /product="NifN"
        /protein_id="AAB91903.1"
        /db_xref="GI:2182684"
        /translation="MWHIRHOSKSTVNPILKSSQPLGAAALVGVQGAIPLPHSGQGC
        TSPALVLCUREPKETIPIQTNDDELATVLGAAHLEEAALNKKRANPLIGIOTTA
        LVRESDPFAKIANIIMTHAEELAGTVLANTPFDGALSEGWAQVADAPLQOITL
        QOAPRGRKATLIRITKPSBQKQKQVAILPWHLTVDGIDQREHREVFQUNPV
        IVPDVGSLDGTVPDRWMPAYGGTSISDIQELGRAVRCIAIGEMRRPAELLQITG
        VYVIVQSLTGLKNDQVPSLLSEICVPAPAKIHRHRSQLODLDGHFHPAGKIIA
        IATPDQYQVATPTGLGABIIISAVTTGESBIIIEKVPARKVQIGDLGLEDLAGGA
        DILLVTHSHGRQAARLEGIPIIRIGPPIFDRLGSHKLTVLYRGTRDLIFEAAITIQAN
        QPAPSLRQIDAMKERNAG"
        /size="6199"
        /genes="nifX"
        /size="6199"
        /genes="nifX"
        /note="nitrogen fixation protein NifX homo.og"

```



```

/codon_start=1
/transl_table=11
/product="NifX"
/protein_id="AA891904.1"
/db_xref="GI:2182685"
/translation="MNSVRLSLVTOEIQPESGALRIATQDMKVLNHFSGAKHP
AVTVDSRGDPFVAVSFDVDSGKPIEKDRITPTVLTGCHLFCRAIGGFS
RAVVSARKHPHIVGPEALQVLSHTOKMLATPPWLRKVLQAGVAKKUPFED
"
6685..7173
/gene="y4v0"
CDS
6685..7173
/gene="y4v0"
/notes="hypothetical 17.7 kd protein, similar to proteins
of other nitrogen-fixing bacteria and to Y4x2"
/codon_start=1
/transl_table=11
/product="y4v0"
/protein_id="AA891905.1"
/db_xref="GI:2182686"
/translations="MKTSTNCTGPDVYKDEALVTPFVKCLARLVRAQTYGSGDRA
SDALLAHPIITBQRRIPIIGDPGVNKAQNFITPAVALSIRATGVAQPIPMI
SHGFGRIVLTGRLVVFSTKVRDHRFGFATRLKLAAGAKLVDDATAIETYPVA
RA"

```

gene

CDS

misc_feature

misc_feature

gene

CDS

gene

CDS

misc_feature

misc_feature

ORIGIN

```

Query Match      60.6%; Score 20; DB 1; Length 10242;
Best Local Similarity 82.1%; Pred. No. 1e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 CTCTTCTCCGAGCCCGTCCGAATAGTGA 31
DB 2986 CTCTTCCAGCAGCCCGTCCGACCTGCTGA 3013

```

RESULT 36

CBRG44D18

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CBRG44D18 42430 bp DNA linear INV 04-NOV-2000
Caenorhabditis briggsae cosmid G44D18, complete sequence.

AC084612 AC084612.1 GI:11095051
HTG.

Caenorhabditis briggsae

Caenorhabditis briggsae

Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

1 (bases 1 to 42430)

The C. briggsae Genome Sequencing Project

Unpublished

2 (bases 1 to 42430)

Wilson, R.

The sequence of C. briggsae cosmid G44D18

Unpublished

3 (bases 1 to 42430)

Waterston, R.

Direct Submission

Submitted (04-NOV-2000) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

Submitted by:

Genome Sequencing Center

Department of Genetics, Washington University,

St. Louis, MO 63110, USA

e-mail: jspietlowatson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES
source

1..42430
/organism="Caenorhabditis briggsae"
/mol_type="genomic DNA"
/strain="Gujarat G16"
/db_xref="taxon:6238"
/clone="G44D18"

ORIGIN

Query Match 60.6%; Score 20; DB 3; Length 42430;
Best Local Similarity 82.1%; Pred. No. 1.2e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

QY 5 TCTTCTCCGAGCCCGTCCGAATAGTGA 32
DB 6131 TATTCTCAGTCCCGTCCGAGTATTGAG 6158

```

RESULT 37

AC084186/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AC084186 53371 bp DNA linear HTG 14-OCT-2000
Homo sapiens clone RP11-702C13, LOW-PASS SEQUENCE SAMPLING.

AC084186 AC084186.1 GI:10801414

HTG; HTGS PHAS80;

Homo sapiens (human)

Homo sapiens

Rukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 53371)

Birren, B., Linton, L., Nuebaum, C. and Lander, E.

Homo sapiens, clone RP11-702C13

Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 53371)
 Birren, B., Linton, L., Nusbaum, C., Lander, B., Abraham, H., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bada, P., Boguslavsky, L.,
 Bouhageltier, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
 DeCarillano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
 Fitzhugh, W., Gage, D., Gilegan, J., Gardyna, S., Ginde, S., Goyette, M.,
 Graham, D., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
 Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K.,
 Lamazares, R., Landers, T., Lechoczky, J., Levine, R., Lieu, C., Liu, G.,
 MacDonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
 McPheeters, R., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
 Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
 O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
 Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
 Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Tirrell, A., Travers, M., Trigglio, J., Vassiliev, R., Vici, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, M.J., Young, G., Zainoun, J.,
 Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (14-OCT-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smith, A.P.A. & Green, P. (1995-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIGR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 Project information
 Center project name: L11454
 Center clone name: 702_C_13

COMMENT

NOTE: This record contains 65 individual
 sequencing reads that have not been assembled into
 contigs. Runs of N are used to separate the reads
 and the order in which they appear is completely
 arbitrary. Low-pass sequence sampling is useful for
 identifying clones that may be gene-rich and allows
 overlap relationships among clones to be deduced.
 However, it should not be assumed that this clone
 will be sequenced to completion. In the event that
 the record is updated, the accession number will
 be preserved.

1 723: contig of 723 bp in length
 724 823: gap of 100 bp
 824 1565: contig of 742 bp in length
 1566 1665: gap of 100 bp
 1666 2414: contig of 749 bp in length
 2415 3251: contig of 737 bp in length
 3252 3351: gap of 100 bp
 3352 4069: contig of 718 bp in length
 4070 4169: gap of 100 bp
 4170 4863: contig of 694 bp in length
 4864 4963: gap of 100 bp
 4964 5669: contig of 706 bp in length
 5670 5769: gap of 100 bp
 5770 6498: contig of 729 bp in length
 6499 6599: gap of 100 bp
 6599 7324: contig of 726 bp in length
 7325 7424: gap of 100 bp
 7425 8152: contig of 728 bp in length
 8153 8252: gap of 100 bp
 8253 8973: contig of 721 bp in length
 8974 9073: gap of 100 bp
 9074 9812: contig of 739 bp in length
 9813 9912: gap of 100 bp
 9913 10651: contig of 739 bp in length
 10652 10751: gap of 100 bp
 10751 11497: contig of 746 bp in length
 11498 11597: gap of 100 bp
 11598 12466: contig of 769 bp in length
 12467 13195: contig of 729 bp in length
 13196 14025: contig of 730 bp in length
 14026 14125: gap of 100 bp
 14126 14864: contig of 739 bp in length
 14865 15696: contig of 732 bp in length
 15697 16521: gap of 100 bp
 16522 16621: contig of 725 bp in length
 16622 17359: contig of 738 bp in length
 17360 18133: gap of 100 bp
 18134 18233: contig of 734 bp in length
 18234 18988: contig of 695 bp in length
 18989 19085: gap of 100 bp
 19086 19905: contig of 717 bp in length
 19906 20628: gap of 100 bp
 20629 21454: contig of 726 bp in length
 21455 22257: contig of 703 bp in length
 22258 23036: gap of 100 bp
 23037 23196: contig of 739 bp in length
 23197 23925: gap of 100 bp
 23926 24742: contig of 723 bp in length
 24743 24842: gap of 100 bp
 24843 25558: contig of 716 bp in length
 25559 26387: gap of 100 bp
 26388 26487: contig of 729 bp in length
 26488 27213: gap of 100 bp
 27214 28047: contig of 734 bp in length
 28048 28862: gap of 100 bp
 28863 29668: contig of 715 bp in length
 29669 30514: gap of 100 bp
 30515 31109: contig of 706 bp in length
 31110 32119: gap of 100 bp
 32120 32953: contig of 722 bp in length
 32954 33775: gap of 100 bp
 33776 34566: contig of 691 bp in length
 34567 35382: gap of 100 bp
 35383 36075: contig of 716 bp in length
 36076 36792: gap of 100 bp
 36793 37715: contig of 593 bp in length
 37716 38540: gap of 100 bp
 38541 39383: contig of 723 bp in length
 39384 40207: gap of 100 bp
 40208 41038: contig of 743 bp in length
 41039 4207: contig of 724 bp in length
 42078 43007: gap of 100 bp
 43008 44038: contig of 731 bp in length

* 41039 41138: gap of 100 bp
 * 41139 41874: contig of 736 bp in length
 * 41875 41974: gap of 100 bp
 * 41975 42699: contig of 725 bp in length
 * 42700 42799: gap of 100 bp
 * 43118 43517: contig of 718 bp in length
 * 43618 44326: contig of 709 bp in length
 * 44327 44426: gap of 100 bp
 * 44427 45136: contig of 730 bp in length
 * 45137 45256: gap of 100 bp
 * 45257 45974: contig of 718 bp in length
 * 45975 46074: gap of 100 bp
 * 46075 46803: contig of 729 bp in length
 * 46804 46903: gap of 100 bp
 * 46904 47625: contig of 722 bp in length
 * 47626 47725: gap of 100 bp
 * 47726 48452: contig of 727 bp in length
 * 48453 48552: gap of 100 bp
 * 48553 49354: contig of 702 bp in length
 * 49355 49354: gap of 100 bp
 * 49355 50084: contig of 730 bp in length
 * 50085 50184: gap of 100 bp
 * 50185 50888: contig of 704 bp in length
 * 50889 50988: gap of 100 bp
 * 50989 51713: contig of 725 bp in length
 * 51714 51813: gap of 100 bp
 * 51814 52552: contig of 739 bp in length
 * 52553 52652: gap of 100 bp
 * 52653 53371: contig of 719 bp in length.

FEATURES

Location/Qualifiers
 1..53371
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="RP11-702C13"
 /clone_lib="RPCI-11 Human Male BAC"

ORIGIN

Query Match 60.6%; Score 20; DB 2; Length 53371;

Best Local Similarity 79.3%; Pred. No. 1.2e+03; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy 5 TCTTCTCGAGCCCGTCAAAATAGTGAGT 33

Db 26584 TCTTCTGTGTCAGGTCGACATAGTGAGT 26556

RESULT 38

AC140870/c 84743 bp DNA linear HTG 05-MAR-2003
 LOCUS Homo sapiens chromosome 5 clone RP11-1132B16, WORKING DRAFT
 DEFINITION SEQUENCE, 5 unordered pieces.

ACCESSION AC140870.1 GI:28849981

VERSION HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ACTIVEPFIN.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 84743)

DOE Joint Genome Institute.

Sequencing of Human Chromosome 5

Unpublished

2 (bases 1 to 84743)

DOE Joint Genome Institute.

Direct Submission

Submitted (05-MAR-2003) Production Sequencing Facility, DOE Joint

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

 Project Information
 Center Project Name: 1729232
 Center clone name: RPCI-11_1132B16

Summary Statistics

Consensus quality: 82347 bases at least Q40
 Consensus quality: 82926 bases at least Q30
 Consensus quality: 83281 bases at least Q20
 Estimated insert size: 175000, agarose-gel estimation
 Estimated insert size: 84543, sum-of-contigs estimation
 Quality coverage: 6.25 in Q20 bases; agarose-gel estimation
 Quality coverage: 12.97 in Q20 bases; sum-of-contigs estimation.
 * NOTES: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1739: contig of 1739 bp in length
 * 1740 1839: gap of unknown length
 * 1840 3234: contig of 1395 bp in length
 * 3235 3334: gap of unknown length
 * 3335 5969: contig of 2635 bp in length
 * 5970 6069: gap of unknown length
 * 6070 12880: contig of 6811 bp in length
 * 12881 12980: gap of unknown length
 * 12981 84743: contig of 71763 bp in length.

FEATURES

Location/Qualifiers
 1..84743
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-1132B16"
 /clone_lib="RPCI human BAC library 11"

ORIGIN

Query Match 60.6%; Score 20; DB 2; Length 84743;

Best Local Similarity 82.1%; Pred. No. 1.2e+03; Indels 0; Gaps 0;

Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 4 CTCCTCTCGAGCCCGTCAAAATAGTGA 31

Db 35233 CTCCTCTCGAGATGATAGATAGTGA 35206

RESULT 39

HS3173N24 85952 bp DNA linear PRI 13-SEP-2001
 LOCUS Human DNA sequence from clone RP11-373N24 on chromosome 6. Contains
 DEFINITION part of a putative novel gene, ESTs, STSs, GSSs and two CpG
 islands, complete sequence.

ACCESSION AL121912.19 GI:8894624

VERSION HTG; CpG island.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 85952)

Parker, A.

Direct Submission

Submitted (30-APR-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: hunquerry@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Jul 1, 2000 this sequence version replaced gi:8346909.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Sm., SMDJ; Swi., SWISSPROT; Tr., TRXBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/c elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> IMPORTANT: This sequence is not the entire insert of clone RP11-373N24. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone RP11-373N24 is at 85952 in this sequence. The true right end of clone RP5-1186N24 is at 100 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-373N24 is from the library RPCI-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES

Source

Location/Qualifiers

1. .85952
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="6"
/clone_lib="RPCI-11.2"

7. .38
/notes="LMC4 repeat: matches 7450. .7479 of consensus"
39. .360
/notes="AluSg repeat: matches 1. .309 of consensus"
361. .635
/notes="LMC4 repeat: matches 7479. .7779 of consensus"
1643. .1717
/notes="ENNA-Ala-GCG repeat: matches 1. .75 of consensus"
1891. .2028
/notes="19 copies 2 mer tt 78% conserved"
2396. .2293
/notes="LTR14 repeat: matches 3. .548 of consensus"
3204. .3488
/notes="AluJo repeat: matches 1. .290 of consensus"
3586. .3808
/notes="LA repeat: matches 4078. .4314 of consensus"
3855. .4097
/notes="LMC5 repeat: matches 6978. .7227 of consensus"
4254. .4422
/notes="LMC5 repeat: matches 7283. .7469 of consensus"
4433. .4637
/notes="AluSg repeat: matches 3. .209 of consensus"
4727. .5020
/notes="AluSg repeat: matches 1. .297 of consensus"
complement(4977. .5480)
/notes="match: GSS: Em:AQ521901"
5517. .5739
/notes="match: GSS: Em:AQ341608"
5517. .5737
/notes="match: GSS: Em:AQ202500"
5521. .5739
/notes="match: GSS: Em:AQ343565"
5536. .5661
/notes="match: GSS: Em:AQ817683"
5759. .5865
/notes="match: GSS: Em:AQ223965"
5876. .6021
/notes="match: GSS: Em:AQ223965"

6512. .6820
/notes="AluY repeat: matches 5. .311 of consensus"
7637. .7958
/notes="match: GSS: Em:AQ831171"
complement(7839. .8442)
/notes="match: GSS: Em:AQ533892"
complement(8150. .8444)
/notes="match: GSS: Em:AQ710263"
8411. .8749
/notes="AluSg repeat: matches 2. .310 of consensus"
8944. .9245
/notes="AluSx repeat: matches 7. .308 of consensus"
9831. .10104
/notes="AluSx repeat: matches 3. .299 of consensus"
complement(10582. .11236)
/notes="match: GSS: Em:AQ586503"
10645. .11098
/notes="CpG island"
/evidence="not_experimental"
11046. .11384
/notes="LTR35 repeat: matches 1. .447 of consensus"
12477. .11807
/notes="AluSx repeat: matches 1. .310 of consensus"
11833. .12055
/notes="AluJb repeat: matches 1. .236 of consensus"
12292. .12947
/notes="PABLA repeat: matches 1. .660 of consensus"
complement(12930. .13351)
/notes="match: GSS: Em:AQ383623"
complement(12931. .13363)
/notes="match: GSS: Em:AQ747522"
13016. .13350
/notes="match: GSS: Em:AQ532682"
13016. .13306
/notes="AluJo repeat: matches 7. .304 of consensus"
13019. .13372
/notes="match: GSS: Em:AQ214591"
13019. .13350
/notes="match: GSS: Em:AQ214568"
complement(13137. .13417)
/notes="match: GSS: Em:AQ236371"
13352. .13399
/notes="match: GSS: Em:AQ585379"
complement(13306. .13840)
/notes="match: GSS: Em:AQ369166"
complement(join(13306. .13370,13754. .13822))
/notes="match: GSS: Em:AQ914771"
complement(13306. .13423)
/notes="match: GSS: Em:AQ2416410"
complement(join(13317. .13415,13748. .13813))
/notes="match: STS: Em:G31524"
13484. .13755
/notes="MER21B repeat: matches 486. .794 of consensus"
13586. .14294
/notes="AluSx repeat: matches 2. .310 of consensus"
14295. .14509
/notes="LTR16A repeat: matches 211. .416 of consensus"
14510. .14786
/notes="AluSp repeat: matches 1. .287 of consensus"
14787. .14994
/notes="LTR16A repeat: matches 1. .211 of consensus"
15272. .15734
/notes="L1MA10 repeat: matches 5169. .5620 of consensus"
15731. .21871
/notes="L1PA3 repeat: matches 15. .6146 of consensus"
21879. .22622
/notes="L1MA10 repeat: matches 5596. .6320 of consensus"
22636. .22935
/notes="AluSg repeat: matches 1. .300 of consensus"
23074. .23104
/notes="MLT2CB repeat: matches 474. .504 of consensus"
23254. .23696
/notes="MLT2CB repeat: matches 10. .461 of consensus"


```
repeat_region /note="MER42 repeat: matches 1. .131 of consensus"
12156 .12448
repeat_region /note="MER42 repeat: matches 211. .504 of consensus"
16738 .16791
repeat_region /note="MER repeat: matches 54. .114 of consensus"
18431 .18468
repeat_region /note="19 copies of 2 mer 87 & conserved"
19958 .21148
repeat_region /note="LI repeat: matches 4197. .5387 of consensus"
21004 .21887
repeat_region /note="LIP2 repeat: matches 1. .891 of consensus"
24234 .25027
repeat_region /note="LIP12 repeat: matches 912. .113 of consensus"
25036 .25335
repeat_region /note="ALUY repeat: matches 301. .2 of consensus"
25338 .25826
repeat_region /note="LI repeat: matches 5356. .4861 of consensus"
27399 .30541
repeat_region /note="LI repeat: matches 4081. .959 of consensus"
30542 .31494
repeat_region /note="LTX5 repeat: matches 969. .14 of consensus"
31503 .32685
repeat_region /note="LI repeat: matches 1207. .6 of consensus"
34179 .34478
repeat_region /note="ALUSP repeat: matches 303. .3 of consensus"
34688 .34792
repeat_region /note="MER repeat: matches 179. .48 of consensus"
35516 .35797
repeat_region /note="ALUSG repeat: matches 303. .28 of consensus"
incomplete repeat
36098 .36401
repeat_region /note="ALUSG repeat: matches 1. .303 of consensus"
36830 .37134
repeat_region /note="MER21B repeat: matches 386. .78 of consensus"
37193 .37492
repeat_region /note="ALUSG repeat: matches 1. .300 of consensus"
37505 .37600
repeat_region /note="16 copies of 6 mer 90 & conserved"
37603 .37668
repeat_region /note="10 copies of 6 mer 85 & conserved"
40195 .40496
repeat_region /note="ALUY repeat: matches 1. .301 of consensus"
40782 .40903
repeat_region /note="2 copies of 61 mer 84 & conserved"
44783 .45080
repeat_region /note="ALUY repeat: matches 292. .1 of consensus"
50516 .50601
repeat_region /note="LIP23 repeat: matches 889. .804 of consensus"
50809 .50925
repeat_region /note="MER30 repeat: matches 5. .116 of consensus"
51044 .51327
repeat_region /note="ALUSX repeat: matches 302. .1 of consensus"
51558 .51682
repeat_region /note="MER6 repeat: matches 865. .738 of consensus"
51785 .52465
repeat_region /note="LIM46 repeat: matches 1028. .294 of consensus"
51787 .52286
repeat_region /note="LIM49 repeat: matches 1020. .529 of consensus"
52492 .53361
repeat_region /note="LI repeat: matches 5358. .4470 of consensus"
53433 .53630
repeat_region /note="ALUSG repeat: matches 298. .103 of consensus"
incomplete repeat
53882 .54095
repeat_region /note="LI repeat: matches 4089. .4311 of consensus"
54106 .54245
repeat_region /note="FLAM C repeat: matches 131. .1 of consensus"
54364 .54459
repeat_region /note="ALUSG repeat: matches 1. .96 of consensus"
incomplete repeat
54510 .54571
repeat_region /note="MER6 repeat: matches 803. .864 of consensus"
55169 .55442
```

```
repeat_region /note="LIM410 repeat: matches 738. .456 of consensus"
55439 .55858
repeat_region /note="MSTA repeat: matches 1. .426 of consensus"
55844 .56236
repeat_region /note="LIM31 repeat: matches 462. .41 of consensus"
56336 .56934
repeat_region /note="LI repeat: matches 5144. .4530 of consensus"
57079 .57188
repeat_region /note="LIP13 repeat: matches 782. .892 of consensus"
57235 .57576
repeat_region /note="LIM32 repeat: matches 770. .421 of consensus"
57777 .58051
repeat_region /note="ALUY repeat: matches 301. .6 of consensus"
58511 .59531
repeat_region /note="LIM43 repeat: matches 1055. .3 of consensus"
59385 .59791
repeat_region /note="LI repeat: matches 5390. .5013 of consensus"
60099 .60215
repeat_region /note="LI repeat: matches 5019. .4903 of consensus"
60234 .60527
repeat_region /note="ALUSX repeat: matches 296. .1 of consensus"
60550 .61291
repeat_region /note="LI repeat: matches 4855. .4064 of consensus"
63556 .63858
repeat_region /note="MER21B repeat: matches 388. .79 of consensus"
```

Query Match 60.68; Score 20; DB 9; Length 118767;
Best Local Similarity 82.14; Pred. No. 1.2e+03;
Matches 23; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 4 CTCTCTCCGAGCGCGTCGCAATAGTGA 31
DB 30751 CTCTCTCCGAGCGTCGCAATAGTGA 30724

Search completed: May 24, 2004, 11:40:56
Job time : 978.689 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 10:22:21 ; Search time 1822.47 seconds
(without alignments)
540.723 Million cell updates/sec

Title: US-10-144-679-1

Perfect score: 33

Sequence: 1 catctcttcgagccggtcgaaatagtgagt 33

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 27513289 seqs, 14911090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthm:*
3: em_estin:*
4: em_estma:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_estc:*
9: gb_est1:*
10: gb_est2:*
11: gb_est3:*
12: gb_est4:*
13: gb_est5:*
14: gb_est6:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pig:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22.4	67.9	517	9	AL886016
C 2	22.4	67.9	758	13	BX773718
C 3	21.6	65.5	727	13	CAL43457
C 4	20.8	63.0	284	12	BM618139

5	20.6	62.4	302	13	BY052208
6	20.6	62.4	354	13	BY081398
7	20.6	62.4	362	13	BY023907
8	20.6	62.4	367	13	BY021474
9	20.6	62.4	373	13	BY025872
10	20.6	62.4	373	13	BY072793
11	20.6	62.4	373	29	CG495817
12	20.6	62.4	375	13	BY010961
13	20.6	62.4	375	13	BY011790
14	20.6	62.4	387	13	BY027371
15	20.6	62.4	391	13	BY773596
16	20.6	62.4	398	13	BY047031
17	20.6	62.4	404	29	CG492941
18	20.6	62.4	415	28	AQ221494
19	20.6	62.4	422	13	BY001675
20	20.6	62.4	426	29	CG492510
21	20.6	62.4	437	10	BB864281
22	20.6	62.4	472	10	BB371463
23	20.6	62.4	489	14	CA576637
24	20.6	62.4	490	12	BY319986
25	20.6	62.4	496	29	CG614436
26	20.6	62.4	502	10	BF023046
27	20.6	62.4	506	10	BP469040
28	20.6	62.4	506	14	CA565719
29	20.6	62.4	509	29	CG531045
30	20.6	62.4	513	29	CG531045
31	20.6	62.4	526	14	CA574401
32	20.6	62.4	527	14	CF007050
33	20.6	62.4	550	13	BO828319
34	20.6	62.4	573	9	AT788228
35	20.6	62.4	575	14	CD554154
36	20.6	62.4	598	14	CA574494
37	20.6	62.4	599	9	AA175402
38	20.6	62.4	604	28	AZ365610
39	20.6	62.4	627	10	BE305737
40	20.6	62.4	638	14	CF913216
41	20.6	62.4	657	10	BB661788
42	20.6	62.4	668	10	BB661788
43	20.6	62.4	671	9	AA869918
44	20.6	62.4	681	13	BY738421
45	20.6	62.4	694	13	BY725617

ALIGNMENTS

RESULT 1

AL886016

LOCUS

DEFINITION

sequence.

AL886016

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL886016 XGC-egg Silurana tropicalis cDNA clone Tegg030m08 5', mRNA
517 bp mRNA linear EST 03-DEC-2003
AL886016.2 GI:38676980
EST.
Silurana tropicalis (western clawed frog)
Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 517)
Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
On Sep 16, 2002 this sequence version replaced gi:22936567.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 200:
TROPICALIS_SEQUENCE_ID: Tegg030m08.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gens Collection (XGC) library
constructed by Aaron M. Zorn.

cdna was oligo dt primed from sug of poly A+ RNA from egg.
 EcorI-NotI cut cDNA was then ligated into pCS107 with EcorI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcorI; Site 2: NotI
 Host: Escherichia coli XL1-blue.

FEATURES

Location/Qualifiers

1..517

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TGG030m08"

/dev_stage="egg"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-egg"

/note="Vector: pCS107; Site 1: EcorI; Site 2: NotI; cDNA

was oligo dt primed from sug of poly A+ RNA from egg.

EcorI-NotI cut cDNA was then ligated into pCS107 with

EcorI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 67.9%; Score 22.4; DB 9; Length 517;
 Best Local Similarity 81.2%; Pred. No. 1.5e+02;
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATCTCTTCCGAGCGCTCGAATAGTAGT 33
 |||||
 Db 220 ATCTCTTCCGAGCGCTCGAATATCAAGAGT 251

RESULT 2

EX773718/c

LOCUS

DEFINITION

EX773718 XGC-egg Silurana tropicalis cDNA clone TB95046n01 3', mRNA

sequence.

EX773718

EX773718.1 GI:39680923

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Silurana tropicalis (western clawed frog)

Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 758)

Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.

Sanger Xenopus tropicalis EST project 2001 (11_2003)

Unpublished (2003)

Contact: Croning MDR

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: trop@sanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TB95046n01.q1x77

Sequencing primer: 77

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.

cdna was oligo dt primed from sug of poly A+ RNA from egg.

EcorI-NotI cut cDNA was then ligated into pCS107 with EcorI at the

5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcorI; Site 2: NotI

Host: Escherichia coli XL1-blue.

FEATURES

Location/Qualifiers

1..758

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TGG046n01"

/dev_stage="egg"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-egg"

/note="Vector: pCS107; Site 1: EcorI; Site 2: NotI; cDNA

was oligo dt primed from sug of poly A+ RNA from egg.

EcorI-NotI cut cDNA was then ligated into pCS107 with

EcorI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 67.9%; Score 22.4; DB 13; Length 758;
 Best Local Similarity 81.2%; Pred. No. 1.5e+02;
 Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ATCTCTTCCGAGCGCTCGAATAGTAGT 33
 |||||
 Db 538 ATCTCTTCCGAGCGCTCGAATATCAAGAGT 507

RESULT 3

LOCUS

DEFINITION

CA143457 727 bp mRNA linear EST 24-SRP-2303

3', mRNA sequence.

CA143457

ACCESSION

VERSION

KEYWORDS

SOURCE

Saccharum officinarum

Saccharum officinarum

Rukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD

clade; Panicoideae; Andropogoneae; Saccharum.

1 (bases 1 to 727)

Vettore, A.L., da Silva, P.R., Kemper, E.L. and Arruda, P.

The libraries that made SUCST

Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

Contact: Arruda P

Centro de Biologia Molecular e Engenharia Genetica

Universidade Estadual de Campinas

Caixa Postal 6010, 13083-970, Campinas SP, Brazil

Tel: 55 19 3788 1137

Fax: 55 19 3788 1089

Email: parruda@unicamp.br

Clone distribution: clone distribution information can be found

through the Brazilian Clone Collection Center (BCCC) at

http://www.bcccenter.fcav.unesp.br

Plate: 035 X96, A column: 07

Seq primer: SP6 Promoter primer.

Location/Qualifiers

1..727

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

/clone="SCQSR2035A07"

/lab_host="DH10B"

/clone_lib="RT2"

/note="Organ: Root tips (0.3cm-long) from adult plants;

Vector: pSPort1; Site 1: SalI; Site 2: NotI; An

unidirectional cDNA library generated from (Root

tips (0.3cm-long) from adult plants). cDNA was prepared

from polyA+ mRNA using Superscript Plusmid System kit

(Invitrogen). The double-strand cDNAs were fractionated

in a sepharose CL-2B 40cm-columns and fragments sizing

between 0.8 and 1.5 kb were directionally cloned into the

vector. Details of each source of RNA and library

construction can be obtained at

http://sucst.lad.ic.unicamp.br/public"

ORIGIN

Query Match 65.5%; Score 21.6; DB 13; Length 727;
 Best Local Similarity 85.7%; Pred. No. 3.4e+02;
 Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TCTTCTCCGAGCGCTCGAATAGTAGT 32
 |||||
 Db 77 TCTTCTCCGAGCGCTCGATATAGTAGT 50

RESULT 4

LOCUS

EX618139

DEFINITION

EX618139 284 bp mRNA linear EST 25-FEB-2002

DEFINITION	170005899386 A.Gam.ad.cdna1 Anopheles gambiae cDNA clone
ACCESSION	BM618139
VERSION	BM618139.1 GI:18916366
KEYWORDS	EST
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae
REFERENCE	1 (bases 1 to 284)
AUTHORS	Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab, R., Collins, P.H., Venter, J.C. and Hoffman, S.L.
TITLE	Celera Anopheles gambiae EST project
JOURNAL	Unpublished (2002)
COMMENT	Contact: Holt R.A. Celera Genomics 45 W. Gude Dr., Rockville, MD 20850, USA Tel: 2404533151 Fax: 2404534580 Email: Holtra@celera.com Plate: MD1004SES row: B column: 12 Seq primer: M13 Reverse.
FEATURES	Location/Qualifiers 1..284 /organism="Anopheles gambiae" /mol_type="mRNA" /strain="RSP-ST (Reduced suso. to Permethrin - std. chromosome)" /db_xref="taxon:71165" /clone="19600447011637" /db_xref="Adult" /dev_stage="DHI0b" /lab_host="DHI0b" /clone_lib="A.Gam.ad.cdna1" /note="vector: pSport1; Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen. CDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."
ORIGIN	Query Match 63.0%; Score 20.8; DB 12; Length 284; Best Local Similarity 78.1%; Pred. No. 6.2e+02; Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0; 1 CATCTCTCTCCGAGCGGTGCGAATAGTGG 32 187 CATGCCATCTCCGAGCGGTGCGAATAGTGG 156
LOCUS	BY052208 302 bp mRNA linear EST 06-DEC-2002
DEFINITION	BY052208 RIKEN full-length enriched, TIB-55 BB88 Mus musculus cDNA clone 1730068L22 5', mRNA sequence.
ACCESSION	BY052208
VERSION	BY052208.1 GI:26157656
KEYWORDS	EST
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	1 (bases 1 to 302)
AUTHORS	Oikazaki, Y., Furuno, M., Kamukawa, T., Adachi, J., Bono, H., Kondo, S., Nakaide, I., Otsu, N., Saito, K., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Cojocari, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Khanapin, A., Matsuda, H., Batalov, S., Beisel, K.M., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, B., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
DEFINITION	170005899386 A.Gam.ad.cdna1 Anopheles gambiae cDNA clone
ACCESSION	BM618139
VERSION	BM618139.1 GI:18916366
KEYWORDS	EST
SOURCE	Anopheles gambiae (African malaria mosquito)
ORGANISM	Anopheles gambiae
REFERENCE	1 (bases 1 to 284)
AUTHORS	Holt, R.A., Lin, J.-J., Murphy, S.D., Evans, C.A., Kraft, C.L., Charlab, R., Collins, P.H., Venter, J.C. and Hoffman, S.L.
TITLE	Celera Anopheles gambiae EST project
JOURNAL	Unpublished (2002)
COMMENT	Contact: Holt R.A. Celera Genomics 45 W. Gude Dr., Rockville, MD 20850, USA Tel: 2404533151 Fax: 2404534580 Email: Holtra@celera.com Plate: MD1004SES row: B column: 12 Seq primer: M13 Reverse.
FEATURES	Location/Qualifiers 1..284 /organism="Anopheles gambiae" /mol_type="mRNA" /strain="RSP-ST (Reduced suso. to Permethrin - std. chromosome)" /db_xref="taxon:71165" /clone="19600447011637" /db_xref="Adult" /dev_stage="DHI0b" /lab_host="DHI0b" /clone_lib="A.Gam.ad.cdna1" /note="vector: pSport1; Site 1: SalI; Site 2: NotI; Whole adult mosquitoes (mixed sex) frozen on liquid nitrogen. CDNA inserts >500 bp cloned directionally into pSport 1. Not 1 site is 3'. Clones available through the Malaria Research and Reference Reagent Resource Center (www.malaria.mr4.org)."
ORIGIN	Query Match 63.0%; Score 20.8; DB 12; Length 284; Best Local Similarity 78.1%; Pred. No. 6.2e+02; Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0; 1 CATCTCTCTCCGAGCGGTGCGAATAGTGG 32 187 CATGCCATCTCCGAGCGGTGCGAATAGTGG 156
LOCUS	BY052208 302 bp mRNA linear EST 06-DEC-2002
DEFINITION	BY052208 RIKEN full-length enriched, TIB-55 BB88 Mus musculus cDNA clone 1730068L22 5', mRNA sequence.
ACCESSION	BY052208
VERSION	BY052208.1 GI:26157656
KEYWORDS	EST
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	1 (bases 1 to 302)
AUTHORS	Oikazaki, Y., Furuno, M., Kamukawa, T., Adachi, J., Bono, H., Kondo, S., Nakaide, I., Otsu, N., Saito, K., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Cojocari, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Khanapin, A., Matsuda, H., Batalov, S., Beisel, K.M., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, B., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,

269 TCTTCTCCGAGCGCGGTGAAGTAGAGA 295

DB BY081398 354 bp mRNA linear EST 07-DEC-2002
 BY081398 RIKEN full-length enriched, 10 days neonate thymus Mus
 musculus cDNA clone X630031N13 5', mRNA sequence.

ACCESSION BY081398
 VERSION BY081398.1 GI:26203084
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 354)

REFERENCES
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriber, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V., Chothia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, N., Hironaka-Kishikawa, T., Konno, H., Nakamura, K., Hayatsu, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL NATURE 420, 563-573 (2002)
 MEDLINE 22346683
 PUBMED 12466851

COMMENT Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-research@riken.go.jp,
 URL: http://genome-research.riken.go.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hironaka, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
 Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genomic Sequences. *Genome* 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
 source Location/Qualifiers
 1..354
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="K630031N13"
 /tissue_type="thymus"
 /dev_stage="10 days neonate"
 /clone_lib="RIKEN full-length enriched, 10 days neonate thymus"

ORIGIN
 Query Match 62.4%; Score 20.6; DB 13; Length 354;
 Best Local Similarity 85.2%; Pred. NO. 7.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCGCGGTGAAGTAGAGA 31
 DB 310 TCTTCTCCGAGCGCGGTGAAGTAGAGA 336

RESULT 7
 BY023907
 LOCUS BY023907
 DEFINITION BY023907 RIKEN full-length enriched, mammary gland RCB-0527
 JY5-MC(B) cDNA Mus musculus cDNA clone G930019H06 5', mRNA sequence.

ACCESSION BY023907.1 GI:26129350
 VERSION BY023907
 KEYWORDS Mus musculus (house mouse)
 SOURCE EST.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 362)

REFERENCES
 AUTHORS Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriber, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V., Chothia, C., Corbani, L.B., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, N., Hironaka-Kishikawa, T., Konno, H., Nakamura, K., Hayatsu, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
MEDLINE
PUBMED
COMMENT

Nature 420, 563-573 (2002)
 22354683
 12466851
 Contact: Yoshihide Hayaishizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9226
 Email: genome-resgsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
 Hayaishizaki, Y. Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES
 source
 1. 362
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G930047105"
 /tissue_type="mammary gland"
 /cell_line="RCB-0527 Jyg-MC(B)"
 /clone_lib="RIKEN full-length enriched, mammary gland
 RCB-0527 Jyg-MC(B) cDNA"

ORIGIN

Query Match 62.4%; Score 20.6; DB 13; Length 362;
 Best Local Similarity 85.2%; Pred. No. 7, 8+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 Oy 5 TCTTCGAGCGCGCGCGGATAGCA 31
 Db 297 TCTTCGAGCGCGCGGATAGGAGA 323

RESULT 8
 BY021474
 LOCUS
 DEFINITION
 BY021474 RIKEN full-length enriched, mammary gland RCB-0526
 JYG-MC(A) cDNA Mus musculus cDNA clone G830047105 5', mRNA
 sequence.

ACCESSION
 BY021474
 VERSION
 BY021474.1 GI:26126917
 EST.

KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Eukaryote; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (base 1 to 367)

REFERENCE
 AUTHORS
 Owaraki, Y., Furuno, M., Kasukawa, T., Aichi, J., Bono, H., Kondo, S.,
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Nature 420, 563-573 (2002)
 22354683
 12466851
 Contact: Yoshihide Hayaishizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9226
 Email: genome-resgsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
 Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
 Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
 Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
 Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and
 Hayaishizaki, Y. Direct Submission
 Computational Analysis of Full-length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

FEATURES
 Location/Qualifiers
 1. 367
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G830047105"
 /tissue_type="mammary gland"
 /cell_line="RCB-0526 Jyg-MC(A)"
 /clone_lib="RIKEN full-length enriched, mammary gland
 RCB-0526 Jyg-MC(A) cDNA"

ORIGIN

Query Match 62.4%; Score 20.6; DB 13; Length 367;
 Best Local Similarity 85.2%; Pred. No. 7.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY 5 TCTTCTCCGAGCGCGTCGAATAGTGA 31
 DB 303 TCTTCTCCGAGCGCGTCGAATAGTGA 327

RESULT 9
 BY025872
 LOCUS
 DEFINITION
 RIKEN full-length enriched, mammary gland RCB-0527 JYG-MC(B) cDNA Mus musculus cDNA clone G930033L4 5', mRNA

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Oikazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Oato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.S., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.P., Forrest, A., Fraser, K.S., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takeda, I., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, T., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, L., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirotsane-Kishikawa, T., Konno, H., Nakamura, M., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)
 22354683
 12466851

CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/

ALIZAWA, K., AKIMURA, T., ARAKAWA, T., ITOH, M., KAWAI, Y., KONNO, H., HIROZANE, T., IMOTO, K., ISHII, Y., ITOH, M., KAWAI, Y., KONNO, H., MIYAZAKI, A., MURATA, M., NAKAMURA, K., NOMURA, K., NUMAZAKI, R., OHNO, M., SAKAI, K., SAKAZUME, N., SASAKI, D., SATO, K., SHIBATA, K., SHIRAKI, T., TAGAMI, M., WAKI, K., WATABIKI, A., MURAMATSU, M. and HAYASHIZAKI, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
 source
 1. 373
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G930033L4"
 /tissue_type="mammary gland"
 /cell_line="RCB-0527 JYG-MC(B)"
 /clone_lib="RIKEN full-length enriched, mammary gland RCB-0527 JYG-MC(B) cDNA"

ORIGIN

Query Match 62.4%; Score 20.6; DB 13; Length 373;
 Best Local Similarity 85.2%; Pred. No. 7.8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

CY 5 TCTTCTCCGAGCGCGTCGAATAGTGA 31
 DB 316 TCTTCTCCGAGCGCGTCGAATAGTGA 342

RESULT 10
 BY0272793
 LOCUS
 DEFINITION
 RIKEN full-length enriched, 16 days embryo heart Mus musculus cDNA clone 1920086F12 5', mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)

REFERENCE
 AUTHORS
 Oikazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Oato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.S., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.P., Forrest, A., Fraser, K.S., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierzki, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takeda, I., Taylor, M.S., Teasdale, R.D., Tomita, K., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, T., Wells, C., Wilming, L.G., Wyshaw-Boris, A., Yanagisawa, M., Yang, L., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirotsane-Kishikawa, T., Konno, H., Nakamura, M., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)
 22354683
 12466851

CONTACT: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsr.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/

ALIZAWA, K., AKIMURA, T., ARAKAWA, T., ITOH, M., KAWAI, Y., KONNO, H., HIROZANE, T., IMOTO, K., ISHII, Y., ITOH, M., KAWAI, Y., KONNO, H., MIYAZAKI, A., MURATA, M., NAKAMURA, K., NOMURA, K., NUMAZAKI, R., OHNO, M., SAKAI, K., SAKAZUME, N., SASAKI, D., SATO, K., SHIBATA, K., SHIRAKI, T., TAGAMI, M., WAKI, K., WATABIKI, A., MURAMATSU, M. and HAYASHIZAKI, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
NATURE 420, 563-573 (2002)

MEDLINE
22354683

PUBMED
12466851

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komori, H., Miyazaki, A., Murata, M., Nakamura, K., Nomura, K., Numazaki, K., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1517-1530 (2000)
RIKEN integrated sequence analysis (RISA) system--184-format sequencing pipeline with 384 multiplexed sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
1..373
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="I920086F12"
/tissue_type="heart"
/dev_stage="16 days embryo"
/clone_id="RIKEN full-length enriched, 16 days embryo heart"

ORIGIN

Query Match 62.4%; Score 20.6; DB 13; Length 373;
Best Local Similarity 85.2%; Pred. No. 7.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCGCGTGGAAATAGTGA 31
|||||
DB 301 TCTTCTCCGAGCGCGTGGAAATAGTGA 327

RESULT 1:1
CG495817
LOCUS
DEFINITION
CG495817 Mus musculus 129SV/EV Mus musculus genomic clones OST35445, genomic survey sequence.
ACCESSION
CG495817
VERSION
CG495817.1 GI:37263500
KEYWORDS
GSS.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches

Conservative

Mismatches

Indels

Gaps

QY

DB

RESULT 12

BY010961

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

TITLE
Analysis of the mouse transcriptome based on functional annotation
Of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683

JOURNAL
MEDLINE
PUBMED

COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsr.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Aizawa, T., Itoh, M., Kawai, J., Konno, H.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.

FEATURES
Location/Qualifiers
1..375
/organism="Mus musculus"
/db_type="taxon:10090"
/db_xref="taxon:10090"
/clone="G3003802"
/tissue="lung"
/clone_lib="RCB-0558 LLC"
cDNA"

ORIGIN
Query Match 62.4% Score 20.6; DB 13, Length 375;
Best Local Similarity 85.2% Pred No. 7.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCAGCCGCTCGAATAGTGA 31
|||||
327 TCTTCTCCAGCCGCTCGAATAGTGA 353

Db 327 TCTTCTCCAGCCGCTCGAATAGTGA 353

RESULT 13
BY011790
LOCUS

DEFINITION
BY011790 RIKEN full-length enriched, lung RCB-0558 LLC cDNA Mus

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse

MUSCULUS
house mouse


```

further details.
Location/Qualifiers
1. .379
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G730029A22"
/tissue_type="lung"
/cell_line="RCS-0558 LLC"
/clone_lib="RIKEN full-length enriched, lung RCS-0558 LLC
cDNA"

ORIGIN
Query Match 62.4%; Score 20.6; DB 13; Length 379;
Best Local Similarity 85.2%; Pred. No. 7.8e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCCGGTCGAATAGTGA 31
Db 316 TCTTCTCCGAGCCGGTCGAATAGTGA 342

RESULT 14
LOCUS BY027371 387 bp mRNA linear EST 06-DEC-2002
DEFINITION Jyg-MC(B) cDNA Mus musculus cDNA clone G930043H08 5', mRNA
sequence.
ACCESSION BY027371 GI:26132814
VERSION BY027371.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 387)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Odate, N., Saito, R., Suzuki, K., Yamana, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuoka, H.,
Batholov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
Cherbas, C., Corbani, L.E., Cousins, S., Dalia, E., Dragani, T.A.,
Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T.,
Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Santelin, A., Schneider, C., Semple, C.A., Secoy, M., Shinada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Waltesred, C., Wang, Y., Watanabe, T.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, L.,
Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirokawa-Kishikawa, T., Konno, H., Kawai, J., Aizawa, K.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, S., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
COMMENT
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

```

```

Tel. 81-45-503-9222
Fax. 81-45-503-9216
Email: genome-resgsc.riken.go.jp/
URL: http://genome-gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,
Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,
Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,
Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,
Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and
Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Yousse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (http://genome-gsc.riken.go.jp) for
further details.
FEATURES
Location/Qualifiers
1. .387
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G930043H08"
/tissue_type="mammary gland"
/cell_line="RCS-0527 Jyg-MC(B)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCS-0527 Jyg-MC(B) cDNA"

ORIGIN
Query Match 62.4%; Score 20.6; DB 13; Length 387;
Best Local Similarity 85.2%; Pred. No. 7.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCCGGTCGAATAGTGA 31
Db 316 TCTTCTCCGAGCCGGTCGAATAGTGA 342

RESULT 15
LOCUS BY773596 391 bp mRNA linear EST 10-DEC-2003
DEFINITION BY773596 RIKEN full-length enriched, 17.5 days embryo whole body
Mus musculus cDNA clone L930074N12 5', mRNA sequence.
ACCESSION BY773596
VERSION BY773596.1 GI:39700234
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 391)
Carninci, P., Waki, K., Shiraki, T., Konno, H., Shibata, K., Itoh, M.,
Aizawa, K., Arakawa, T., Ishii, Y., Sasaki, D., Bono, H., Kondo, S.,
Sugahara, Y., Saito, R., Sato, N., Fukuda, S., Sato, K., Watabiki, A.,
Hirozane-Kishikawa, T., Nakamura, M., Shibata, Y., Yasunishi, A.,
Kikuchi, N., Yoshiki, A., Kusakabe, M., Gustincich, S., Beisel, K.,
Pavan, W., Aizawa, K., Nakagawa, A., Held, W.A., Iwata, H., Kono, T.,
Nakauchi, H., Lyons, P., Wells, C., Hume, D.A., Pagioli, M.,
Hensch, T.K., Brinkmeier, M., Camper, S., Hirota, J., Mombaerts, P.,
Muramatsu, M., Okazaki, Y., Kawai, J. and Hayashizaki, Y.
Tagging a complex transcriptome: the construction of the mouse
full-length cDNA encyclopedia

```

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 13 (68), 1273-1289 (2003)
22703353
12619125
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome-gsc.riken.jp/>) for
further details.

FEATURES
source

Location/Qualifiers
1. 391
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="L930074N12"
/tissue_type="whole body"
/dev_stage="17.5 days embryo"
/clone_lib="RIKEN full-length enriched, 17.5 days embryo
whole body"

ORIGIN

Query Match 62.4%; Score 20.6; DB 13; Length 391;
Best Local Similarity 85.2%; Pred. No. 7.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCGAGCCGGTCGAATAGTGA 31
|||||
DB 307 TCTTCTCGAGCCGGTCGTAAGTAGA 333
|||||

RESULT 16
BY047031
LOCUS

DEFINITION
BY047031 RIKEN full-length enriched, TIB-55 B888 Mus musculus cDNA
clone I730044J17 5', mRNA sequence.

ACCESSION
BY047031

VERSION
BY047031.1

KEYWORDS
BY047031.1

SOURCE
Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 398)

REFERENCE

1. Oikawa, Y., Furumoto, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, T., Otsu, N., Saito, R., Suzuki, H., Yamana, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gajjar, T., Baldarelli, R., Hill, D.P., Bult, C.,
Hume, D.A., Quackenbush, J., Schriml, L.M., Manapin, A., Matsuda, H.,
Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Flachner, C.P., Forrest, A., Fraser, K.S., Gassterland, T.,
Gariboldi, M., Glass, C., Godik, A., Gough, J., Grimmond, S.,
Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, S.D., Kanai, A.,
Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lemhard, B., Lyons, P.A., Maglott, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, T., Pavan, N.J., Perlea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Read, J.C., Read, D.J., Reid, J., Ring, B.Z., Ringwald, M.,
Santana, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,

Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, S., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome-gsc.riken.jp/>) for
further details.

FEATURES
source

Location/Qualifiers
1. 398
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="I730044J17"
/cell_line="TIB-55 B888"
/clone_lib="RIKEN full-length enriched, TIB-55 B888"

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2): 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site (<http://genome-gsc.riken.go.jp/>) for
further details.

Query Match 62.4%; Score 20.6; DB 13; Length 398;
Best Local Similarity 85.2%; Pred. No. 7.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

FEATURES
source

Location/Qualifiers
1. 398
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="I730044J17"
/cell_line="TIB-55 B888"
/clone_lib="RIKEN full-length enriched, TIB-55 B888"

ORIGIN

Query Match 62.4%; Score 20.6; DB 13; Length 398;
Best Local Similarity 85.2%; Pred. No. 7.9e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY

5 TCTTCTCGAGCCGGTCGAATAGTGA 31

DB

269 TCTTCTCGAGCCGGTCGTAAGTAGA 295

RESULT 17

CC492941

LOCUS

DEFINITION

ACCESSION

VERSION

CC492941
OST31011 Mus musculus 1295v/Bv Mus musculus genomic clone OST31011,
genomic survey sequence.
CC492941
CC492941.1 GI:37257647

KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

Zambrowicz, R.P., Abuln, A., Ramirez-Solis, R., Richter, L.J., Piggott, J., Beltrando, R.H., Buxton, E.C., Edwards, J., Finch, R.A., Priddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jiang, C., Kay, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z., Sparks, M.J., Van Sighemorst, I., Vogel, P., Walke, M., Xu, N., Zhu, Q., Person, C. and Sands, A.T.

TITLE
 Waki kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention

JOURNAL
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

CONTACT: Zambrowicz BP

OmniBank
 Lexicon Genetics Incorporated
 4000 Research Forest Drive, The Woodlands, TX 77381, USA
 Email: materials@lexgen.com

Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al [Nature. 1998 Apr 9;392(6676):1608-11]

Class: Gene Trap.

Location/Qualifiers
 1. .404
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="129SV/Ev"
 /db_xref="taxon:10090"
 /clone="OST31011"
 /cell_type="embryonic stem cell"
 /clone_lib="Mus musculus 129SV/Ev"

ORIGIN

Query Match 62.4%; Score 20.6; DB 29; Length 404;
 Best Local Similarity 85.2%; Pred. No. 7.9e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TCTCTCCGAGCGCGTGAATAGTA 31
Db 168 TCTCTCCGAGCGCGTGAATAGTA 194

RESULT 18
LOCUS
DEFINITION
 HS 2004_B2_G06_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2004 Col=12 Row=N, genomic survey sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)

REFERENCE
AUTHORS

Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Purlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

TITLE
 Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome.

JOURNAL
MEDLINE
PUBLISHED
COMMENT

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

93480589
 10449764

Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu

Sequence Tagged Connector
 Plate: 2004 row: N column: 12
 Class: BAC ends
 High quality sequence stop: 415.
 Location/Qualifiers
 1. .415
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="plate=2004 Col=12 Row=N"
 /sex="male"
 /clone_lib="CIT Approved Human Genomic Sperm Library D"
 /notes="Organ: sperm; Vector: pBlOBAcl1; BAC Clones in E-Coli DH108"

ORIGIN

Query Match 62.4%; Score 20.6; DB 28; Length 415;
 Best Local Similarity 85.2%; Pred. No. 8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 TCTCTCCGAGCGCGTGAATAGT 29
Db 343 TCTCTCCGAGCGCGTGAATAGT 369

RESULT 19
LOCUS
DEFINITION
 BY001675 RIKEN full-length enriched, 15 days embryo head Mus musculus cDNA clone 4022405A17 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Mus musculus (house mouse)

REFERENCE
AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojobori, T., Balderelli, R., Hill, D.P., Bult, C., Hume, D.A., Chackebush, J., Schimi, L.M., Kanapin, A., Masuda, H., Bencalov, S., Balcel, K.W., Blake, J.A., Bradt, D., Brusci, V., Chochia, C., Corbani, L.S., Cousins, S., Daille, S., Dragani, T.A., Fletcher, C.P., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Glasi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, S.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltale, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Nunata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramakrishnan, S., Ravasi, T., Read, J.C., Read, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, N.S., Teasdale, R.D., Tomita, K., Verdaro, R., Wagner, L., Wahlstedt, C., Wang, Y., Watarabe, Y., Wells, C., Wilming, I.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kuroki, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imorani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Tsunishi, A., Yoshino, M., Waterson, R., Lander, S.S., Rogers, J., Birney, R. and Hayashizaki, Y.

TITLE
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL
MEDLINE
PUBLISHED
COMMENT

Nature 420, 563-573 (2002)

12466851

Contact: Yoshitake Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

DEFINITION	CG9730401 Mus musculus 129Sv/Ev Mus musculus genomic clone CG9730401, genomic survey sequence.
ACCESSION	CG492510
VERSION	CG492510.1 GI:37256796
KEYWORDS	G88.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 426)
AUTHORS	Zambrowicz, B.P., Abuin, A., Ramirez-Solis, R., Richter, L.J., Piggott, J., Beltrando-Rello, H., Buxton, B.C., Edwards, J., Finch, R.A., Friddle, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jiang, C., Key, B.W., Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D., Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrnick, J., Shi, Z.-Z., Sparks, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N., Zhu, Q., Person, C. and Sands, A.T.
TITLE	Wnk1 kinase deficiency lowers blood pressure in mice: a gene-trap screen to identify potential targets for therapeutic intervention
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)
COMMENT	Contact: Zambrowicz BP OmniBank Lexicon Genetics Incorporated 4000 Research Forest Drive, The Woodlands, TX 77381, USA Email: material@lexgen.com
FEATURES	Gene trap sequence tag generated by 3' RACE from mouse ES cells as described in Zambrowicz et al (Nature, 1998 Apr 9;392(6767):608-11) Class: Gene trap. Location/Qualifiers 1. 426
SOURCE	

1

23; Conservative	62.4%;	Score 20.6;
1.1 Similarity	85.2%;	Pred. No. 8;
		0; Mismatch

QY	5	TCTTCTCCGAGCCCGGTGGAAATAGTGA	31
DB	168	TCTTCTCCGAGCCCGGTGAAGTAGAGA	194
RESULT 21			
BB864281			
LOCUS			
DEFINITION			
ACCESSION	BB864281	mRNA	linear EST 09-JUL-2003
VERSION	BB864281	musculus full-length enriched, RCB-0464 Meth-A CDNA MAB	
KEYWORDS	BB864281.1	musculus cDNA clone G430042015 5', mRNA sequence.	
SOURCE	GI:17110491		
ORGANISM	Mus musculus (house mouse)		
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

Authors

Akamura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T., Hayase, N., Hiramoto, K., Hiraoaka, T., Hirozane, T., Imotani, K., Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sorabe, Y., Suzuki, H., Tagawa, A., Takebaashi, F., Takaku-Akahira, S.,

TITLE RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al. 2001)
JOURNAL Unpublished (2001)

COMMENT

Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-resgsc.riken.go.jp,
 URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayashizaki, Y., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.,
 and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,
 Sugahara, Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

8 mouse tissues.

Location/Qualifiers
 1. 437
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="BALB/c"
 /db_xref="taxon:10090"
 /clone="G430042015"
 /sex="female"
 /tissue_type="sarcoma"
 /cell_line="RCB-0464 Meth-A"
 /clone_lib="RIKEN full-length enriched, RCB-0464 Meth-A
 cDNA"

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 437;
 Best Local Similarity 85.2%; Pred. NO. 8e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QV 5 TCTTCTCCGAGCCGCTCGAAATAGTGA 31
 |||||
 DB 318 TCTTCTCCGAGCCGCTCGAAATAGTGA 344
 |||||

RESULT 22

BE371463
 LOCUS 472 bp mRNA linear EST 21-JUL-2000
 DEFINITION 601223163F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE3592119 5',
 mRNA sequence.

ACCESSION BE371463

VERSION BE371463.1 GI:9316826

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 472)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM8762 row: m column: 16

High quality sequence stop: 428.

FEATURES

source

Location/Qualifiers
 1. 472
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE3592119"
 /tissue_type="spontaneous tumor, metastatic to mammary.
 Stem cell origin."
 /lab_host="DH108"
 /clone_lib="NCI CGAP Lu29"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 472;
 Best Local Similarity 85.2%; Pred. NO. 8.1e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QV 5 TCTTCTCCGAGCCGCTCGAAATAGTGA 31
 |||||

DB 278 TCTTCTCCGAGCCGCTCGAAATAGTGA 304
 |||||

RESULT 23

CAS76637

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CAS76637 489 bp mRNA linear EST 19-NOV-2002
 X0647G11-5N NIA Mouse Hematopoietic Stem Cell (lin-/c-Kit-/Sca-1+)
 cDNA Library (Long) Mus musculus cDNA clone NIA:X0647G11
 IMAGE:30073330 5', mRNA sequence.

CAS76637

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 489)

Piao, Y., Margul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,

Taub, D., Longo, D.L., Keller, J. and Xie, M.S.H.

Systematic Analyses of NIA Mouse Hematopoietic Stem Cell

(lin-/c-Kit-/Sca-1+) cDNA Library (Long)

Unpublished (2001)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institutes on Aging/National Institutes of Health

333 Casell drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@gsun.gcr.nia.nih.gov

Plate: X0647 row: G column: 11

Seq primer: M13 Reverse

High quality sequence stop: 489

POLYA:No.

FEATURES

source

Location/Qualifiers
 1. 489
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6NCr"
 /db_xref="niaEST:X0647G11-5N"
 /db_xref="taxon:10090"
 /clone="NIA:X0647G11 IMAGE:30073330"
 /tissue_type="Hematopoietic Stem Cell
 (lin-/c-Kit-/Sca-1+)"
 /dev_stage="Age approx. 10 weeks old"
 /lab_host="DH108"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell

(Lin/-c-Kit/-Sca-1+) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (<http://igun.grc.nia.nih.gov/cDNA/>). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001)). [PMID: 11541991]. Total RNAs were
 obtained from Drs. Dennis Taub, Dan Longo (National
 Institute on Aging, USA), Jonathan Keller (National Cancer
 Institute, USA). Double-stranded cDNAs were synthesized
 with an Oligo(dT) primer (Invitrogen):
 5'-pGACTAGTTCAGATCGAGCGCCCTTCTTTTCTTTT-3' from
 1.1 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lona-linker L1-SalI, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 5x Taq polymerase (Takara) with a primer SalI-3. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.2 kb. The library was constructed
 by Yulan Piao (NIA)."

ORIGIN

Query Match 62.4%; Score 20.6; DB 14; Length 489;
 Best Local Similarity 85.2%; Pred. No. 8.2e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 3;

Qy 5 TCTTCTCCGAGCGCGTCCGAATAGTGA 31
 |||||
 Db 220 TCTTCTCCGAGCGCGTGAAGTAGAGA 246

RESULT 24

B1319986 490 bp mRNA linear EST 21-JUL-2001
 LOCUS le46001.y1 Kaestner ngm3 wc Mus musculus cDNA 5 similar to
 SW:GABB MOUSE Q00420 GA BINDING PROTEIN BETA-1 CHAIN 1, mRNA
 sequence.

ACCESSION B1319986.1 GI:14998672

VERSION B1319986

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 490)

AUTHORS Melton, D., Brown, J., Kenty, G., Bermitt, A., Lee, C., Kaestner, K.,

Hillier, L., Marra, M., Pape, D., Wyllie, T., Martin, J., Blisstein, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J.,

Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, J.,

Williams, T., Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohpc.harvard.edu

Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,

2000) Library was constructed by Catherine Lee DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Dr. Marie Searce

(msearce@mail.med.upenn.edu)

Seq primer: -40RP from Glibco

High quality sequence stop: 350.

FEATURES

source

Location/Qualifiers

1..490
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="129/Sv x CD1"
 /db_xref="taxon:10090"
 /seq_start="p.c. 14.5"
 /seq_end="p.c. 14.5"
 /clone_lib="Kastner ngm3 wt"
 /note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site 1:
 NotI; Site 2: SalI; The library was prepared by
 Catherine S. Lee and has not been published. The pancreas
 was obtained from Gerard Gradwohl (PNAS 97 P1607-1611,
 2000). The cDNA's were prepared with an oligo containing a
 NotI site, and SalI linkers were added to the ends. The
 inserts were cut with NotI before being cloned into the
 NotI-SalI sites in the vectors. This is one of two
 libraries, ngm3 wt and ngm3 -/-.. The wt library is in
 pSPORT1, T7 promoter is 5'."

ORIGIN

Query Match 62.4%; Score 20.6; DB 12; Length 490;
 Best Local Similarity 85.2%; Pred. No. 8.2e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 5 TCTTCTCCGAGCGCGTCCGAATAGTGA 31
 |||||
 Db 211 TCTTCTCCGAGCGCGTGAAGTAGAGA 237

RESULT 25

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436 496 bp DNA linear GSS 02-OCT-2003

QST302692 Mus musculus 129Sv/Sv Mus musculus genomic clone

QST302692, genomic survey sequence.

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436.1 GI:37438285

QST302692 Mus musculus 129Sv/Sv Mus musculus genomic clone

QST302692, genomic survey sequence.

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CG614436

LOCUS

DEFINITION

CG614436

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

ACCESSION BP469040
 VERSION BP469040.1 GI:11538223
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 REFERENCE 1 (bases 1 to 506)
 AUTHORS Bonaldo, M.P., Lennon, G. and Soares, M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6051 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20832-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: m35@mail.nih.gov
 cDNA Library Preparation: M.B. Soares Lab Clone distribution:
 Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
 should be noted that Bento Soares is generating a small number of
 additional specialized non-redundant arrays of BMAP cDNAs whose
 availability will be considered under appropriate and limited
 collaborative arrangements
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..506
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /dev_stage="27-32 days"
 /clone="UI-M-BH3-att-h-12-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M 84"
 /note="Vector: p7735-Pac (Pharmacia) with a modified
 polylinker, Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP M 84 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH BMAP M 84,
 NIH BMAP M 83.3, NIH BMAP M 83.2, NIH BMAP M 83.1,
 NIH BMAP M 82, NIH BMAP M 81. The subtracted library
 (NIH BMAP M 84) was constructed as follows: PCR amplified
 cDNA inserts from NIH BMAP M 83.3, NIH BMAP M 83.2, and
 NIH BMAP M 83.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH BMAP M 83.3, NIH BMAP M 83.2, and NIH BMAP M 83.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH BMAP M 84 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)."

FEATURES

source

1..506
 Location/Qualifiers
 1..506
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /dev_stage="27-32 days"
 /clone="UI-M-BH3-att-h-12-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH BMAP M 84"
 /note="Vector: p7735-Pac (Pharmacia) with a modified
 polylinker, Site 1: Not I; Site 2: Eco RI; The
 NIH BMAP M 84 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH BMAP M 84,
 NIH BMAP M 83.3, NIH BMAP M 83.2, NIH BMAP M 83.1,
 NIH BMAP M 82, NIH BMAP M 81. The subtracted library
 (NIH BMAP M 84) was constructed as follows: PCR amplified
 cDNA inserts from NIH BMAP M 83.3, NIH BMAP M 83.2, and
 NIH BMAP M 83.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH BMAP M 83.3, NIH BMAP M 83.2, and NIH BMAP M 83.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH BMAP M 84 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)."

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 506;
 Best Local Similarity 85.2%; Pred. No. 8.2e+02;
 Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCCGCGTAAGTAGAGA 31
 |||||
 DB 149 TCTTCTCCGAGCCGCGTAAGTAGAGA 175

RESULT 29

CG530185

LOCUS

DEFINITION

CG530185 Mus musculus 129SV/Ev Mus musculus genomic clone

ACCESSION

CG530185

VERSION

CG530185.1 GI:37316757

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 509)

AUTHORS

Browicz, B.P., Abulin, R., Ramirez-Solis, R., Richter, L.J.,

Piggott, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,

Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,

Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,

Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,

Spark, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,

Zhu, Q., Person, C. and Sande, A.T.

TITLE

Wnt1 kinase deficiency lowers blood pressure in mice: a gene-trap

screen to identify potential targets for therapeutic intervention

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 100 (24), 14109-14114 (2003)

COMMENT

Contact: Zambrowicz BP

OmiBank

Lexicon Genetics Incorporated

4000 Research Forest Drive, The Woodlands, TX 77381, USA

Email: materials@lexgen.com

Gene trap sequence tag generated by 3' RACE from mouse ES cells as

described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)

Class: Gene trap.

Location/Qualifiers

1..509

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="129SV/Ev"

/db_xref="taxon:10090"

/clone="OST11625"

/cell_type="embryonic stem cell"

/clone_lib="Mus musculus 129SV/Ev"

ORIGIN

Query Match 62.4%; Score 20.6; DB 29; Length 509;

Best Local Similarity 85.2%; Pred. No. 8.2e+02;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCCGCGTAAGTAGAGA 31

|||

DB 175 TCTTCTCCGAGCCGCGTAAGTAGAGA 201

RESULT 30

CG531045

LOCUS

DEFINITION

CG531045 Mus musculus 129SV/Ev Mus musculus genomic clone

ACCESSION

CG531045

VERSION

CG531045.1 GI:37317617

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 513)

AUTHORS

Zambrowicz, B.P., Abulin, R., Ramirez-Solis, R., Richter, L.J.,

Piggott, J., BeltrandelRio, H., Buxton, E.C., Edwards, J., Finch, R.A.,

Fridde, C.J., Gupta, A., Hansen, G., Hu, Y., Huang, W., Jaing, C.,

Key, B.W. Jr., Kipp, P., Kohlhauff, B., Ma, Z.-Q., Markesich, D.,

Payne, R., Potter, D.G., Qian, N., Shaw, J., Schrick, J., Shi, Z.-Z.,

Spark, M.J., Van Sligtenhorst, I., Vogel, P., Walke, W., Xu, N.,

Zhu, Q., Person, C. and Sande, A.T.

TITLE Mki kinase deficiency lowers blood pressure in mice: a gene-trap
screen to identify potential targets for therapeutic intervention
JOURNAL Proc Natl Acad Sci U S A. 100 (24), 14109-14114 (2003)
COMMENT Contact: Zambrowicz BP

CanisBank
Lexicon Genetics Incorporated
4000 Research Forest Drive, The Woodlands, TX 77381, USA
Email: materials@lexgen.com
Gene trap sequence tag generated by 3' RACE from mouse ES cells as
described in Zambrowicz et al (Nature. 1998 Apr 9;392(6676):608-11)
Class: Gene Trap.

FEATURES

source

Location/Qualifiers
1..513
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129SV/EV"
/db_xref="taxon:10090"
/clone="G0113328"
/cell_type="embryonic stem cell"
/clone_lib="Mus musculus 129SV/EV"

ORIGIN

Query Match 62.4%; Score 20.6; DB 29; Length 513;
Best Local Similarity 85.2%; Pred. No. 8.2e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCCGGTCAATAGTGA 31
DB 175 TCTTCTCCGAGCCGGTCAATAGTGA 201

RESULT 31

CA574401

LOCUS K0615G08-SN NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
DEFINITION cDNA Library (Long) Mus musculus cDNA clone NIA:K0615G08
IMAGE:30070255 5', mRNA sequence.

CA574401

CA574401.1 GI:25119102

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 526)

Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.

Systematic Analyses of NIA Mouse Hematopoietic Stem Cell

(Lin-/c-Kit-/Sca-1+) cDNA Library (Long)

Unpublished (2001)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: K0615 row: G column: 08

Seq primer: M13 Reverse

High quality sequence stop: 526

POLYA=NO.

Location/Qualifiers

source

1..526
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="nia:EST:K0615G08-5N"
/db_xref="taxon:10090"
/clone="NIA:K0615G08 IMAGE:30070255"
/tissue_type="Hematopoietic Stem Cell"
/lin-/c-Kit-/Sca-1+)
/dev_stage="Age approx. 10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell"
/clone_lib="NIA Mouse Hematopoietic Stem Cell"
/lin-/c-Kit-/Sca-1+ cDNA Library (Long)

/note=Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National Cancer
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer (Invitrogen):
5'-pGACTGATGTCGAGCCGGTCAATAGTGA-3' from
1.1 ug of total RNA, treated with 1% DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lona-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.2 Kb. The library was constructed
by Yulan Piao (NIA)."

ORIGIN

Query Match 62.4%; Score 20.6; DB 14; Length 526;
Best Local Similarity 85.2%; Pred. No. 8.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCCGGTCAATAGTGA 31
DB 220 TCTTCTCCGAGCCGGTCAATAGTGA 246

RESULT 32

CF907050

LOCUS

DEFINITION

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

Mus musculus

```

/dev stage="9-15C cells"
/lab host="DK108"
/clone lib="NIA Mouse Mesenchymal Stem Cell cDNA Library
(long 1)"
/notes="vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://agum.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2003)). [PMID: 11541199]. Total
RNAs were obtained from Dr. Akhihiro Umezawa (Keio
University School of Medicine, Japan). Double-stranded
cDNAs were synthesized with an Oligo(dT) primer
(Invitrogen):
5'-TGACTGATCTAGATCGGAGCGCCGCTTTTCTTTT-3' from
2.2 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to long-linker L1-Sal4, purified by phenol/chloroform, and
separated from free linkers by centrifugation 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.5 kb. The library was
constructed by Yulan Piao."

```

ORIGIN

```

Query Match 62.4%; Score 20.6; DB 14; Length 527;
Best Local Similarity 85.2%; Pred. No. 8.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 5 TCTCTCCGAGCGCGTGCAGATAGTGA 31
|||||
DB 246 TCTCTCCGAGCGCGTGCAGATAGTGA 272
|||||

```

RESULT 33

```

BQ838319/c
LOCUS
DEFINITION
WHE2909 B02_C0325 Wheat aluminum-stressed root tip cDNA library
Triticum aestivum cDNA clone WHE2909_B02_C03, mRNA sequence.
ACCESSION
BQ838319
VERSION
BQ838319.1 GI:22142637
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 550)
Anderson, O.D., Chao, S., Chin, A., Close, T.J., Gustafson, J.P.,
Lazo, G.R., Rausch, C.J., Ross, K., Seaton, C.L. and Wilson, C.
The structure and function of the expressed portion of the wheat
genomes - Aluminum-stressed root tip cDNA library
Unpublished (2001)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@wv.ars.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: SK primer.

```

FEATURES

```

source
1..550
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="BHL146"

```

```

/db_xref="taxon:4565"
/clone="WHE2909_B02_C03"
/tissue_type="Root tip at 1.0 to 1.5 mm stage"
/dev stage="Seedling"
/lab host="E. coli SOLR"
/clone lib="Wheat aluminum-stressed root tip cDNA library"
/notes="vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI; Plants were grown under
hydroponic conditions, root tips were excised and snap
frozen. Total RNA was prepared at University of
Missouri (Ross, Gustafson) Poly(A) RNA was purified, a
cDNA library was made, and the cDNA clones were in vivo
excised to give phagescript SK- phagemids in the T3 Close
lab (Chin and Close) at the University of California,
Riverside. Plasmid DNA preparations and DNA sequencing
were performed in the OD Anderson lab (all other
authors)."

```

ORIGIN

```

Query Match 62.4%; Score 20.6; DB 13; Length 550;
Best Local Similarity 85.2%; Pred. No. 8.3e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 7 TTCTCCGAGCGCGTGCAGATAGTGA 33
|||||
DB 223 TGCTCCGAGCGCGTGCAGATAGTGA 197
|||||

```

RESULT 34

```

A1788228
LOCUS
DEFINITION
U119A07.Y1 Sugano mouse embryo mewa Mus musculus cDNA clone
IMAGE:2087988 5' similar to gb:M74516 Mouse GA binding protein
(MOORE)., mRNA sequence.
ACCESSION
A1788228
VERSION
A1788228.1 GI:5335944
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 573)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
Person, B., Swallow, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R.,
Ritter, S., Kohn, S., Shin, J., Jackson, J., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The Mouse Genome Sequencing Consortium
Other ESTs: u119a07.x1
Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1803
Fax: 314 286 1810
Email: mouse@wustl.wustl.edu
This clone is available royalty-free through LNCX, contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:995672

```

```

Seq primer: custom primer used
High quality sequence stop: 502.

```

FEATURES

```

source
1..573
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL"
/db_xref="taxon:10090"
/clone="IMAGE:2087988"
/dev stage="embryo, 14 dpc"
/lab host="DH10B"
/clone lib="Sugano mouse embryo mewa"
/notes="vector: PM8185-PL3; Site 1: DraIII (CACTGTGTG);
Site 2: DraIII (CACCATGTC); 1st strand cDNA was primed

```


11: 1553-1558 (2001). (PMID: 115441991). Total RNAs were obtained from Drs. Dennis Taub, Dan Longo (National Institute on Aging, USA), Jonathan Keller (National Cancer Institute, USA). Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen]: 5'-pGACGATCTAGACCGGAGCGCCGCTTTTCTTTT-3' from 1.1 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lene-linker LL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 2.2 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 62.4%; Score 20.6; DB 14; Length 598;
Best Local Similarity 85.2%; Pred. No. 8.4e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCGCGTCCGAAATAGTGA 31
|||||
DB 220 TCTTCTCCGAGCGCGTGTAGTAGAGA 246
|||||

RESULT 37

AA175402 599 bp mRNA linear EST 16-FEB-1997
US80A06.r1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE 617842
5' similar to gb:U74517 Mouse GA binding protein (MOUSE), mRNA
sequence.

AA175402.1 GI:1756541

AA175402.1 GI:1756541

EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 599)

Marta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,

Schellenberg, K., Steptoe, M., Tan, P., Underwood, K., Moore, B.,

Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and

Waterston, A.

The WashU-HMI Mouse EST Project

Unpublished (1996)

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LNC; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 489.

Location/Qualifiers

1. .599

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="IMAGE:617842"

/sex="male"

/tissue_type="Spleen"

/dev_stage="4 weeks"

FEATURES

source

/lab host="DH10B"
/clone lib="Soares mouse 3NDMS"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TCTTACCACTGAGTCGAGCGCGCGCTTTTCTTTT-3']
3); double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M. Patricia Bonaldo."

ORIGIN

Query Match 62.4%; Score 20.6; DB 9; Length 599;
Best Local Similarity 85.2%; Pred. No. 8.4e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 TCTTCTCCGAGCGCGTCCGAAATAGTGA 31
|||||
DB 296 TCTTCTCCGAGCGCGTGTAGTAGAGA 322
|||||

RESULT 38

AA165610 604 bp DNA linear GSS 02-OCT-2000
M0112N19P Mouse 10kb plasmid UUC1M library Mus musculus genomic
clone UUC1M0112N19 P, genomic survey sequence.

AA165610

AA165610.1 GI:10479310

GSS.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 604)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, G., Pedersen, T.,

Rally, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von

Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: daunegenetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0112 row: A column: 19

Seq primer: CGTTGTAACCGCGCCAGT

Class: plasmid ends

High quality sequence stop: 604.

Location/Qualifiers

1. .604

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUC1M0112N19"

/sex="Male"

/lab host="E. Coli strain XL10-Gold, T1-resistant, P-"

/clone lib="Mouse 10kb plasmid UUC1M library"

/note="Vector: pMD42w, Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (gi1473214|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 62.4%; Score 20.6; DB 28; Length 604;
Best Local Similarity 85.2%; Pred. No. 8.5e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 5 TCTTCTCCGAGCCGGTGGTAAGTAGAGA 31
|||||

Db 222 TCTTCTCCGAGCCGGTGGTAAGTAGAGA 248
|||||

RESULT 39

BE305737
LOCUS 627 bp mRNA linear EST 26-OCT-2000
DEFINITION F0110245371 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3494869 5',
mRNA sequence.

ACCESSION BE305737

VERSION BE305737.1 GI:9157887

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 627)

Author: NTH-MGC http://mgc.nci.nih.gov/.

Title: National Institutes of Health, Mammalian Gene Collection (MGC)

Journal: Unpublished (1999)

Comment: Contact: Robert Strausberg, Ph.D.

Email: c9apbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LEM8544 row: i column: 14

High quality sequence stop: 543.

FEATURES

source

1..627

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Czech II"

/db_xref="taxon:10090"

/clone="IMAGE:3494869"

/tissue_type="spontaneous tumor, metastatic to mammary."

Stem cell origin.

/lab_host="DH10B"

/clone_lib="NCI CGAP Lu29"

/notes="Organ: lung; Vector: PCMV-SPORE6; Site: 1: SalI;

Site: 2: NotI; Cloned unidirectionally. Primer: oligo dr.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NTH"

ORIGIN

Query Match 62.4%; Score 20.6; DB 10; Length 627;
Best Local Similarity 85.2%; Pred. No. 8.5e+02;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Cy 5 TCTTCTCCGAGCCGGTGGTAAGTAGAGA 31
|||||

Db 287 TCTTCTCCGAGCCGGTGGTAAGTAGAGA 313
|||||

RESULT 40

LOCUS CF913216

DEFINITION A0645C02-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/sca-1+)

CDNA Library (Long 1) Mus musculus cDNA clone NIA:A0645C02

IMAGE:30751897 5', mRNA sequence.

ACCESSION CF913216

VERSION CF913216.1 GI:38184420

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 638)

Author: Piao Y., KO, N.T., Lim, M.K. and KO, M.S.H.

Title: Construction of long-transcript enriched cDNA libraries from

submicrogram amounts of total RNAs by a universal PCR amplification

method

Genome Res. 11 (9), 1553-1558 (2001)

21429098

11541199

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@gaun.grc.nia.nih.gov

Plate: A0645 row: C column: 02

Seq primer: M13 Reverse

High quality sequence stop: 638

POLFA=NO.

FEATURES

source

1..638

Location/Qualifiers

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6NCR"

/db_xref="nia:EST:A0645C02-5"

/db_xref="taxon:10090"

/clone="NIA:A0645C02 IMAGE:30751897"

/dev_stage="Age -10 weeks old"

/lab_host="DH10B"

/clone_lib="NIA Mouse Hematopoietic Stem Cell

(Lin-/c-Kit-/Sca-1+)

/notes="Vector: PCMV-SPORE6 (Invitrogen); Site: 1: SalI;

Site: 2: NotI; Mouse cDNA project by the Laboratory of

Genetics, National Institute on Aging (NIA), Intramural

Research Program, NIH (http://gaun.grc.nia.nih.gov/cDNA).

This is a long-transcript enriched cDNA library (Ref.

Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]. Total

RNAs were obtained from Drs. Dennis Taub, Dan Longo

(National Institute on Aging, USA), Jonathan Keller

(National Cancer Institute, USA), Double-stranded cDNAs

were synthesized with an oligo(dT) primer [Invitrogen];

5'-pGACTATCTCAGCGAGCGCCGCTTTTCTTTT-3' from

1.1 ug of total RNA, treated with T4 DNA polymerase, and

purified by ethanol-precipitation. The cDNAs were ligated

to Loxe-linker IL-Sal4, purified by phenol/chloroform, and

separated from free linkers by Centricon 100. Then, the

cDNAs were amplified by long-range high fidelity PCR using

Ex Taq polymerase (Takara) with a primer Sal4-S. The

products were purified by phenol/chloroform and Centricon

100. The cDNAs were digested with SalI and NotI enzymes

and cloned into SalI/NotI site of PCMV-SPORE6 plasmid

vector. The DH10B *E. coli* host was transformed with the

ligation mixture by the standard chemical method. The

average insert size is about 2.2 kb. The library was

constructed by Yulan Piao."

ORIGIN

Query Match 62.4%; Score 20.6; DB 14; Length 638;
Best Local Similarity 85.2%; Pred. No. 8.5e+02;

Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 5 TCTTCTCCGAGCCGTCGAAATAGTGA 31
 |||||
 Db 220 TCTTCTCCGAGCCGTCGAAATAGTGA 246
 |||||

Search completed: May 24, 2004, 12:29:49
 Job time : 1826.47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

QM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 10:35:46 ; Search time 30.9434 Seconds
(without alignments)
358.688 Million cell updates/sec

Title: US-10-144-679-2

Perfect score: 20

Sequence: 1 actcactataggagagatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365419

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgm2_6/ptodata/2/ina/5A COMB.seq.*
2: /cgm2_6/ptodata/2/ina/5B COMB.seq.*
3: /cgm2_6/ptodata/2/ina/6A COMB.seq.*
4: /cgm2_6/ptodata/2/ina/6B COMB.seq.*
5: /cgm2_6/ptodata/2/ina/PTCUS COMB.seq.*
6: /cgm2_6/ptodata/2/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	30	1	US-08-472-194A-9
C 2	20	100.0	30	3	Sequence 9, Appli
C 3	20	100.0	30	3	Sequence 9, Appli
C 4	20	100.0	30	4	US-08-849-567A-9
5	20	100.0	38	4	Sequence 9, Appli
6	20	100.0	38	4	US-08-849-069-8
7	20	100.0	38	4	Sequence 8, Appli
8	20	100.0	38	4	US-08-849-069-53
9	20	100.0	38	4	Sequence 53, Appli
10	20	100.0	43	1	Sequence 11, Appli
11	20	100.0	43	1	Sequence 11, Appli
12	20	100.0	43	3	Sequence 11, Appli
13	20	100.0	68	3	Sequence 11, Appli
14	20	100.0	69	4	Sequence 23, Appli
15	19	95.0	87	4	Sequence 32, Appli
16	19	95.0	87	4	Sequence 32, Appli
17	19	95.0	87	4	Sequence 32, Appli
18	19	95.0	87	4	Sequence 32, Appli
19	18	90.0	107	1	Sequence 13, Appli
20	18	90.0	107	1	Sequence 13, Appli
21	18	90.0	107	1	Sequence 13, Appli
22	17.4	87.0	54	3	Sequence 13, Appli
23	17.4	87.0	60	3	Sequence 13, Appli
24	17.4	87.0	60	3	Sequence 13, Appli
25	17.4	87.0	67	3	Sequence 13, Appli
26	17.4	87.0	67	3	Sequence 13, Appli
27	17	85.0	19	1	Sequence 7, Appli

28	17	85.0	19	1	US-08-472-194A-12	Sequence 12, Appli
29	17	85.0	19	4	US-08-849-567A-7	Sequence 7, Appli
30	17	85.0	19	4	US-08-849-567A-12	Sequence 12, Appli
31	16.8	84.0	40	2	US-08-549-211-38	Sequence 38, Appli
32	16.8	84.0	43	1	US-09-015-876-23	Sequence 23, Appli
33	16.8	84.0	44	2	US-08-441-887A-24	Sequence 24, Appli
34	16.8	84.0	44	3	US-08-544-381B-121	Sequence 121, Appli
35	16.8	84.0	48	4	US-09-710-200-4	Sequence 4, Appli
36	16.8	84.0	48	4	US-09-975-408-4	Sequence 4, Appli
37	16.8	84.0	49	4	US-09-600-770A-5	Sequence 5, Appli
38	16.8	84.0	50	3	US-09-290-577-16	Sequence 16, Appli
39	16.8	84.0	50	4	US-09-290-452-16	Sequence 16, Appli
40	16.8	84.0	50	4	US-09-290-338-16	Sequence 16, Appli
41	16.8	84.0	50	4	US-09-493-491-32	Sequence 32, Appli
42	16.8	84.0	50	4	US-09-290-000-16	Sequence 16, Appli
43	16.8	84.0	50	4	US-09-493-491A-35	Sequence 35, Appli
44	16.8	84.0	50	4	US-09-954-594A-16	Sequence 16, Appli
45	16.8	84.0	50	4	US-09-530-095B-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-08-472-194A-9/c
Sequence 9, Application US/08472194A
Patent No. 5807718
GENERAL INFORMATION:
APPLICANT: Joyce, Gerald P
TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSER: The Scripps Research Institute, Office of
ADDRESSER: Patent Counsel
STREET: 10550 No. 5807718th Torrey Pines Road, "PC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,194A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/349,023
FILING DATE: 02-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI 463.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-784-2937
TELEFAX: 619-784-9399
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-472-194A-9

Query Match 100.0%; Score 20; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ACTCATTATAGGAGAGATG 20

;; PRIOR FILING DATE: 1999-06-18
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: MS-DOS
;; SEQ ID NO 57
;; LENGTH: 38
;; TYPE: DNA
;; ORGANISM: artificial sequence
;; FEATURE:
;; OTHER INFORMATION: primer
US-09-849-069-57

Query Match 100.0%; Score 20; DB 4; Length 38;
Best Local Similarity 100.0%; Pred. No. 0.39; Indels 0;
Matches 20; Conservative 0; Mismatches 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 14 ACTCACTATAGGAGAGATG 33

RESULT 7
US-08-472-194A-11
;; Sequence 11, Application US/08472194A
;; Patent No. 5807718
;; GENERAL INFORMATION:
;; APPLICANT: Joyce, Gerald P
;; APPLICANT: Breaker, Ronald R
;; TITLE OF INVENTION: ENZYMAIC DNA MOLECULES
;; NUMBER OF SEQUENCES: 40
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: The Scripps Research Institute, Office of
;; ADDRESS: Patent Counsel
;; STREET: 10550 No. 5807718th Torrey Pines Road, TPC 8
;; CITY: La Jolla
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 92037

;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/472,194A
;; FILING DATE: 07-JUN-1995
;; CLASSIFICATION: 536
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/349,023
;; FILING DATE: 02-DEC-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Fitting, Thomas
;; REGISTRATION NUMBER: 34,163
;; REFERENCE/DOCKET NUMBER: ISRI 463.1
;; TELEPHONE: 619-784-2937
;; TELEFAX: 619-784-9399
;; INFORMATION FOR SEQ ID NO: 11:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 43 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-472-194A-11

Query Match 100.0%; Score 20; DB 1; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 8
US-09-262-142-11
;; Sequence 11, Application US/09262142B
;; Patent No. 6110462
;; GENERAL INFORMATION:
;; APPLICANT: Barbas, Carlos P.
;; APPLICANT: Joyce, Gerald
;; APPLICANT: Santoro, Stephen W.
;; APPLICANT: Kandamam, Sakthivel
;; TITLE OF INVENTION: ENZYMAIC DNA MOLECULES THAT CONTAIN MODIFIED
;; FILE REFERENCE: SCR21448
;; CURRENT APPLICATION NUMBER: US/09/262,142B
;; CURRENT FILING DATE: 1999-03-03
;; NUMBER OF SEQ ID NOS: 58
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 11
;; LENGTH: 43
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE: Description of Artificial Sequence: primer
US-09-262-142-11

Query Match 100.0%; Score 20; DB 3; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 9
US-09-262-142-13
;; Sequence 13, Application US/09262142B
;; Patent No. 6110462
;; GENERAL INFORMATION:
;; APPLICANT: Barbas, Carlos P.
;; APPLICANT: Joyce, Gerald
;; APPLICANT: Santoro, Stephen W.
;; APPLICANT: Kandamam, Sakthivel
;; TITLE OF INVENTION: ENZYMAIC DNA MOLECULES THAT CONTAIN MODIFIED
;; FILE REFERENCE: SCR21448
;; CURRENT APPLICATION NUMBER: US/09/262,142B
;; CURRENT FILING DATE: 1999-03-03
;; NUMBER OF SEQ ID NOS: 58
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 13
;; LENGTH: 43
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE: Description of Artificial Sequence: substrate
US-09-262-142-13

Query Match 100.0%; Score 20; DB 3; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 10
US-08-849-567A-11

/ Sequence 11, Application US/08849567A
 / Patent No. 6326174
 / GENERAL INFORMATION:
 / APPLICANT: Joyce, Gerald P.
 / TITLE OF INVENTION: ENZYMIC DNA MOLECULES
 / FILE REFERENCE: SCRL9438
 / CURRENT APPLICATION NUMBER: US/08/849,567A
 / PRIOR FILING DATE: 1997-08-25
 / PRIOR APPLICATION NUMBER: PC7/US95/15580
 / PRIOR FILING DATE: 1995-12-01
 / PRIOR APPLICATION NUMBER: 08/472,194
 / PRIOR FILING DATE: 1995-06-07
 / PRIOR APPLICATION NUMBER: 08/349,023
 / PRIOR FILING DATE: 1994-12-02
 / NUMBER OF SEQ ID NOS: 101
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 11
 / LENGTH: 43
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Description of Artificial Sequence: primer
 / US-08-849-567A-11

Query Match 100.0%; Score 20; DB 4; Length 43;
 Best Local Similarity 100.0%; Pred. No. 0.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGATG 20
 DB 19 ACTCACTATAGGAAGATG 38

RESULT 11
 US-09-262-142-23
 / Sequence 23, Application US/09262142B
 / Patent No. 6110462
 / GENERAL INFORMATION:
 / APPLICANT: Barbas, Carlos P.
 / APPLICANT: Joyce, Gerald
 / APPLICANT: Santoro, Stephen W.
 / APPLICANT: Khandasamy, Sakthivel
 / TITLE OF INVENTION: ENZYMIC DNA MOLECULES THAT CONTAIN MODIFIED
 / FILE REFERENCE: SCRL21448
 / CURRENT APPLICATION NUMBER: US/09/262,142B
 / CURRENT FILING DATE: 1999-03-03
 / NUMBER OF SEQ ID NOS: 59
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 23
 / LENGTH: 68
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Description of Artificial Sequence: oligomer
 / NAME/KEY: misc_feature
 / LOCATION: (49)
 / OTHER INFORMATION: n = g, a, c or t
 / US-09-262-142-23

Query Match 100.0%; Score 20; DB 3; Length 68;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGATG 20
 DB 19 ACTCACTATAGGAAGATG 38

RESULT 12
 US-09-849-069-33

/ Sequence 33, Application US/09849069
 / Patent No. 6630306
 / GENERAL INFORMATION:
 / APPLICANT: Ronald R. Breaker
 / TITLE OF INVENTION: Bioreactive Allosteric Polynucleotides
 / FILE REFERENCE: OCR-794.CIP
 / CURRENT APPLICATION NUMBER: US/09/849,069
 / CURRENT FILING DATE: 2001-05-07
 / PRIOR APPLICATION NUMBER: US 09/331,809
 / PRIOR FILING DATE: 1999-06-18
 / NUMBER OF SEQ ID NOS: 57
 / SOFTWARE: MS-DOS
 / SEQ ID NO 33
 / LENGTH: 69
 / TYPE: DNA
 / ORGANISM: artificial sequence
 / FEATURE:
 / OTHER INFORMATION: self-cleaving DNA
 / US-09-849-069-33

Query Match 100.0%; Score 20; DB 4; Length 69;
 Best Local Similarity 100.0%; Pred. No. 0.42;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGATG 20
 DB 14 ACTCACTATAGGAAGATG 33

RESULT 13
 US-09-849-069-32
 / Sequence 32, Application US/09849069
 / Patent No. 6630306
 / GENERAL INFORMATION:
 / APPLICANT: Ronald R. Breaker
 / TITLE OF INVENTION: Bioreactive Allosteric Polynucleotides
 / FILE REFERENCE: OCR-794.CIP
 / CURRENT APPLICATION NUMBER: US/09/849,069
 / CURRENT FILING DATE: 2001-05-07
 / PRIOR APPLICATION NUMBER: US 09/331,809
 / PRIOR FILING DATE: 1999-06-18
 / NUMBER OF SEQ ID NOS: 57
 / SOFTWARE: MS-DOS
 / SEQ ID NO 32
 / LENGTH: 87
 / TYPE: DNA
 / ORGANISM: artificial sequence
 / FEATURE:
 / OTHER INFORMATION: C1 variant DNA
 / US-09-849-069-32

Query Match 100.0%; Score 20; DB 4; Length 87;
 Best Local Similarity 100.0%; Pred. No. 0.43;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGATG 20
 DB 14 ACTCACTATAGGAAGATG 33

RESULT 14
 US-08-472-194A-13
 / Sequence 13, Application US/08472-94A
 / Patent No. 5807718
 / GENERAL INFORMATION:
 / APPLICANT: Joyce, Gerald P.
 / APPLICANT: Breaker, Ronald R.
 / TITLE OF INVENTION: ENZYMIC DNA MOLECULES
 / NUMBER OF SEQUENCES: 40
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: The Scripps Research Institute, Office of
 / ADDRESSEE: Patent Counsel
 / STREET: 10550 No. 5807718th Torrey Pines Road, TPC 8


```

/ APPLICATION NUMBER: US 08/349,023
/ FILING DATE: 02-DEC-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fitting, Thomas
/ REGISTRATION NUMBER: 34,163
/ REFERENCE/DOCKET NUMBER: TSRI 463.1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 619-784-2937
/ TELEFAX: 619-784-9399
/ INFORMATION FOR SEQ ID NO: 23:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 107 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: misc difference
/ LOCATION: replace(28, "A")
/ OTHER INFORMATION: /standard name= "ADENOSINE
/ OTHER INFORMATION: RIBONUCLEOTIDE"
/ OTHER INFORMATION: /label= 2A
US-08-472-194A-23

```

```

Query Match          95.0%; Score 19; DB 1; Length 107;
Best Local Similarity 95.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

```

```

RESULT 18
US-08-849-567A-23
/ Sequence 23, Application US/08849567A
/ Patent No. 6326174
/ GENERAL INFORMATION:
/ APPLICANT: Joyce, Gerald F.
/ TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
/ FILE REFERENCE: SCR19438
/ CURRENT APPLICATION NUMBER: US/08/849,567A
/ CURRENT FILING DATE: 1997-08-25
/ PRIOR APPLICATION NUMBER: PCT/US95/15580
/ PRIOR FILING DATE: 1995-12-01
/ PRIOR APPLICATION NUMBER: 08/472,194
/ PRIOR FILING DATE: 1995-06-07
/ PRIOR APPLICATION NUMBER: 08/349,023
/ PRIOR FILING DATE: 1994-12-02
/ NUMBER OF SEQ ID NOS: 101
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 23
/ LENGTH: 107
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Combined DNA/RNA Molecule: The n
/ OTHER INFORMATION: at position 28 is adenosine ribonucleotide.
/ OTHER INFORMATION: Description of Artificial Sequence: oligomer
/ NAME/KEY: misc feature
/ LOCATION: (49)..(88)
/ OTHER INFORMATION: n is an equimolare mixture of G, A, T and C
US-08-849-567A-23

```

```

Query Match          95.0%; Score 19; DB 4; Length 107;
Best Local Similarity 95.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

```

```

RESULT 19
US-09-262-142-2
/ Sequence 2, Application US/09262142B
/ Patent No. 6110462
/ GENERAL INFORMATION:
/ APPLICANT: Barbas, Carlos F.
/ APPLICANT: Joyce, Gerald
/ APPLICANT: Santoro, Stephen W.
/ APPLICANT: Kandassamy, Sakthivel
/ TITLE OF INVENTION: ENZYMATIC DNA MOLECULES THAT CONTAIN MODIFIED
/ FILE REFERENCE: SCR21448
/ CURRENT APPLICATION NUMBER: US/09/262,142B
/ CURRENT FILING DATE: 1999-03-03
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 19
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: termination
/ OTHER INFORMATION: sequence
/ NAME/KEY: modified_base
/ LOCATION: (8)
/ OTHER INFORMATION: adenosine ribonucleotide
US-09-262-142-2

```

```

Query Match          90.0%; Score 18; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 TCACATATAGGAGAGATG 20
DB 1 TCACATATAGGAGAGATG 18

```

```

RESULT 20
US-09-262-142-7
/ Sequence 7, Application US/09262142B
/ Patent No. 6110462
/ GENERAL INFORMATION:
/ APPLICANT: Barbas, Carlos F.
/ APPLICANT: Joyce, Gerald
/ APPLICANT: Santoro, Stephen W.
/ APPLICANT: Kandassamy, Sakthivel
/ TITLE OF INVENTION: ENZYMATIC DNA MOLECULES THAT CONTAIN MODIFIED
/ FILE REFERENCE: SCR21448
/ CURRENT APPLICATION NUMBER: US/09/262,142B
/ CURRENT FILING DATE: 1999-03-03
/ NUMBER OF SEQ ID NOS: 58
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 7
/ LENGTH: 19
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: primer
/ NAME/KEY: modified_base
/ LOCATION: (8)
/ OTHER INFORMATION: adenosine ribonucleotide
US-09-262-142-7

```

```

Query Match          90.0%; Score 18; DB 3; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 TCACATATAGGAGAGATG 20
DB 1 TCACATATAGGAGAGATG 18

```

Db 1 TCACTATAGGAGAGATG 18

RESULT 21

US-09-262-142-12

Sequence 12, Application US/09262142B

Patent No. 6110462

GENERAL INFORMATION:

APPLICANT: Barbas, Carlos P.

APPLICANT: Joyce, Gerald

APPLICANT: Santoro, Stephen W.

APPLICANT: Kandasamy, Sakthivel

TITLE OF INVENTION: ENZYMATIC DNA MOLECULES THAT CONTAIN MODIFIED

FILE REFERENCE: SCR21448

CURRENT APPLICATION NUMBER: US/09/262.142B

CURRENT FILING DATE: 1999-03-03

NUMBER OF SEQ ID NOS: 58

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 12

LENGTH: 19

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURES:

OTHER INFORMATION: Description of Artificial Sequence: substrate

NAME/KEY: modified_base

LOCATION: (8)

OTHER INFORMATION: adenosine ribonucleotide

US-09-262-142-12

Query Match 90.0%; Score 18; DB 3; Length 19;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCACTATAGGAGAGATG 20

Db 1 TCACTATAGGAGAGATG 18

RESULT 22

US-09-262-142-12

Sequence 2, Application US/08928465

Patent No. 6204024

GENERAL INFORMATION:

APPLICANT: Romano, Joseph

APPLICANT: Lee, Run Mi

TITLE OF INVENTION: CCR5 RNA Transcription Based

OPERATING SYSTEM: PC-DOS/MS-DOS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Akzo No. 6204024el Patent Department

STREET: 1300 Piccard Drive

CITY: Rockville

STATE: Maryland

COUNTRY: US

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/928,465

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gormley, Mary B.

REGISTRATION NUMBER: 34,409

TELECOMMUNICATION INFORMATION:

TELEPHONE: 301-948-7400

TELEFAX: 301-948-9751

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 54 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: not relevant

MOLECULE TYPE: other nucleic acid

DESCRIPTION: /desc = "DNA Oligonucleotide"

HYPOTHETICAL: NO

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..25

OTHER INFORMATION: /label= T7 RNA Polymera

US-08-928-465-2

Query Match 87.0%; Score 17.4; DB 3; Length 54;

Best Local Similarity 94.7%; Pred. No. 7.9;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGAT 19

Db 13 ACTCACTATAGGAGAGAT 31

RESULT 23

US-09-660-877-9

Sequence 9, Application US/09660877

Patent No. 6251639

GENERAL INFORMATION:

APPLICANT: Kuhn, N.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR LINEAR

FILE REFERENCE: 492692000100

CURRENT APPLICATION NUMBER: US/09/660,877

CURRENT FILING DATE: 2000-09-13

NUMBER OF SEQ ID NOS: 22

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 60

TYPE: DNA

ORGANISM: Synthetic primer

FEATURE:

OTHER INFORMATION: IA012b

US-09-660-877-9

Query Match 87.0%; Score 17.4; DB 3; Length 60;

Best Local Similarity 94.7%; Pred. No. 8;

Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGAT 19

Db 8 ACTCACTATAGGAGAGAT 26

RESULT 24

US-09-893-191B-13

Sequence 13, Application US/09893191B

Patent No. 6686156

GENERAL INFORMATION:

APPLICANT: Kuhn, Nurieth

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR

FILE REFERENCE: 492692000600

CURRENT APPLICATION NUMBER: US/09/893,191B

CURRENT FILING DATE: 2001-06-26

PRIOR APPLICATION NUMBER: 60/213,908

PRIOR FILING DATE: 2000-06-26

PRIOR APPLICATION NUMBER: 60/217,748

PRIOR FILING DATE: 2001-03-21

NUMBER OF SEQ ID NOS: 21

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13

LENGTH: 60

TYPE: DNA

;; ORGANISM: Artificial Sequence
 ;; FEATURE:
 ; OTHER INFORMATION: Synthetic Primer: IA012b
 US-09-893-191B-13

Query Match 87.0%; Score 17.4; DB 4; Length 60;
 Best Local Similarity 94.7%; Pred. No. 8;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGAT 19
 DB 8 ACTCACTATAGGAGAGAT 26

RESULT 25
 US-09-660-877-8
 ; Sequence 8, Application US/09660877
 ; Patent No. 6251639
 ; GENERAL INFORMATION:
 ; APPLICANT: Kuhn, N.
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR LINEAR
 ; FILE REFERENCE: 49262000100
 ; CURRENT APPLICATION NUMBER: US/09/660,877
 ; CURRENT FILING DATE: 2000-09-13
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 67
 ; TYPE: DNA
 ; ORGANISM: Synthetic primer
 ; FEATURE:
 ; OTHER INFORMATION: IA012
 US-09-660-877-8

Query Match 87.0%; Score 17.4; DB 3; Length 67;
 Best Local Similarity 94.7%; Pred. No. 8.1;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGAT 19
 DB 15 ACTCACTATAGGAGAGAT 33

RESULT 26
 US-09-893-191B-11
 ; Sequence 11, Application US/09893191B
 ; Patent No. 6686156
 ; GENERAL INFORMATION:
 ; APPLICANT: Kuhn, Nurith
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
 ; FILE REFERENCE: 49262000600
 ; CURRENT APPLICATION NUMBER: US/09/893,191B
 ; CURRENT FILING DATE: 2001-06-26
 ; PRIOR FILING DATE: 2000-06-26
 ; PRIOR APPLICATION NUMBER: 60/213,908
 ; PRIOR FILING DATE: 2000-06-26
 ; PRIOR APPLICATION NUMBER: 60/277,748
 ; PRIOR FILING DATE: 2001-03-21
 ; NUMBER OF SEQ ID NOS: 21
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 67
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Synthetic Primer: IA012
 US-09-893-191B-11

Query Match 87.0%; Score 17.4; DB 4; Length 67;
 Best Local Similarity 94.7%; Pred. No. 8.1;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGAT 19
 DB 15 ACTCACTATAGGAGAGAT 33

RESULT 27
 US-08-472-194A-7
 ; Sequence 7, Application US/08472194A
 ; Patent No. 5807718
 ; GENERAL INFORMATION:
 ; APPLICANT: Joyce, Gerald F
 ; APPLICANT: Breaker, Ronald R
 ; TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
 ; NUMBER OF SEQUENCES: 40
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: The Scripps Research Institute, Office of
 ; ADDRESSER: Patent Counsel
 ; STREET: 10550 No. 5807718th Torrey Pines Road, TPC 8
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/472,194A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/349,023
 ; FILING DATE: 02-DEC-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fitting, Thomas
 ; REGISTRATION NUMBER: 34,163
 ; REFERENCE/DOCKET NUMBER: TSRI 463.1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-784-2937
 ; TELEFAX: 619-784-9399
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 19 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: misc difference
 ; LOCATION: replace(8,"")
 ; OTHER INFORMATION: /standard_name="ADENOSINE"
 ; OTHER INFORMATION: RIBONUCLEOTIDE"
 US-08-472-194A-7

Query Match 85.0%; Score 17; DB 1; Length 19;
 Best Local Similarity 94.4%; Pred. No. 11;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 TCACATATAGGAGAGATG 20
 DB 1 TCACATATAGGAGAGATG 18

RESULT 28
 US-08-472-194A-12
 ; Sequence 12, Application US/08472194A
 ; Patent No. 5807718
 ; GENERAL INFORMATION:
 ; APPLICANT: Joyce, Gerald F
 ; APPLICANT: Breaker, Ronald R
 ; TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
 ; NUMBER OF SEQUENCES: 40

1 CORRESPONDENCE ADDRESS:
2 ADDRESSEE: The Scripps Research Institute, Office of
3 ADDRESSEE: Patent Counsel
4 STREET: 10550 No. 5807718th Torrey Pines Road, TPC 8
5 CITY: La Jolla
6 STATE: CA
7 COUNTRY: USA
8 ZIP: 92037
9 COMPUTER READABLE FORM:
10 MEDIUM TYPE: Floppy disk
11 COMPUTER: IBM PC compatible
12 OPERATING SYSTEM: PC-DOS/MS-DOS
13 SOFTWARE: Patent Release #1.0, Version #1.25
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/472,194A
16 FILING DATE: 07-JUN-1995
17 CLASSIFICATION: 536
18 PRIOR APPLICATION DATA:
19 APPLICATION NUMBER: US 08/349,023
20 FILING DATE: 02-DEC-1994
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Fitting, Thomas
23 REGISTRATION NUMBER: 34,163
24 REFERENCE/DOCKET NUMBER: TSP1 463.1
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: 619-784-2937
27 TELEFAX: 619-784-9199
28 INFORMATION FOR SEQ ID NO: 12:
29 SEQUENCE CHARACTERISTICS:
30 LENGTH: 19 base pairs
31 TYPE: nucleic acid
32 STRANDNESS: single
33 TOPOLOGY: linear
34 MOLECULE TYPE: DNA (genomic)
35 FEATURES:
36 NAME/KEY: misc difference
37 LOCATION: replace(8, "A")
38 OTHER INFORMATION: /standard name= "ADENOSINE
39 OTHER INFORMATION: RIBONUCLEOTIDE"
40 US-08-472-194A-12

Query Match 85.0%; Score 17; DB 1; Length 19;
Best Local Similarity 94.4%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 3 TCACTATAGGAGAGATG 20
Db 1 TCACTATAGGAGAGATG 18

RESULT 29
US-08-449-567A-7
Sequence 7, Application US/08849567A
Patent No. 6326174
GENERAL INFORMATION:
APPLICANT: Joyce, Gerald F.
APPLICANT: Joyce, Gerald R.
TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
FILE REFERENCE: SCR1943S
CURRENT APPLICATION NUMBER: US/08/849,567A
PRIOR FILING DATE: 1997-08-25
PRIOR APPLICATION NUMBER: PCT/US95/15580
PRIOR FILING DATE: 1995-12-01
PRIOR APPLICATION NUMBER: 08/472,194
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 08/349,023
PRIOR FILING DATE: 1994-12-02
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 7
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence

1 FEATURES:
2 OTHER INFORMATION: Description of Combined DNA/RNA Molecule: The n
3 OTHER INFORMATION: at position 8 is adenosine ribonucleotide.
4 OTHER INFORMATION: Description of Artificial Sequence: substrate
5 US-08-849-567A-7

Query Match 85.0%; Score 17; DB 4; Length 19;
Best Local Similarity 94.4%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 3 TCACTATAGGAGAGATG 20
Db 1 TCACTATAGGAGAGATG 18

RESULT 30
US-08-849-567A-12
Sequence 12, Application US/08849567A
Patent No. 6326174
GENERAL INFORMATION:
APPLICANT: Joyce, Gerald F.
APPLICANT: Joyce, Ronald R.
TITLE OF INVENTION: ENZYMATIC DNA MOLECULES
FILE REFERENCE: SCR1943S
CURRENT APPLICATION NUMBER: US/08/849,567A
PRIOR FILING DATE: 1997-08-25
CURRENT APPLICATION NUMBER: PCT/US95/15580
PRIOR FILING DATE: 1995-12-01
PRIOR APPLICATION NUMBER: 08/472,194
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 08/349,023
PRIOR FILING DATE: 1994-12-02
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 12
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Combined DNA/RNA Molecule: The n
OTHER INFORMATION: at position 8 is adenosine ribonucleotide.
OTHER INFORMATION: Description of Artificial Sequence: substrate
US-08-849-567A-12

Query Match 85.0%; Score 17; DB 4; Length 19;
Best Local Similarity 94.4%; Pred. No. 11;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 3 TCACTATAGGAGAGATG 20
Db 1 TCACTATAGGAGAGATG 18

RESULT 31
US-08-549-211-38
Sequence 38, Application US/08549211
Patent No. 5874260
GENERAL INFORMATION:
APPLICANT: CLUZAT, Philippe
APPLICANT: GUILLOU-BONNICI, Françoise
APPLICANT: LEVASSEUR, Pierre
APPLICANT: MAILLET, Françoise
TITLE OF INVENTION: OLIGONUCLEOTIDE WHICH CAN BE USED AS
PRIOR FILING DATE: 1995-12-01
PRIOR APPLICATION NUMBER: 08/472,194
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: 08/349,023
PRIOR FILING DATE: 1994-12-02
NUMBER OF SEQUENCES: 42
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oliff & Berridge
STREET: 700 South Washington Street, Suite 300
CITY: Alexandria
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22314

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: US/08/549,211
FILING DATE: 27-OCT-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Herridge, William P.
REGISTRATION NUMBER: 30,024
REFERENCE/DOCKET NUMBER: WFB 36705
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6400
TELEFAX: (703) 836-2787
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-549-211-38

Query Match 84.0%; Score 16.8; DB 2; Length 40;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Cy 1 ACTCCTATAGGAGAGATG 20
Db 8 ACTCCTATAGGAGAGATG 27

RESULT 32.
US-09-015-876-23
Sequence 23, Application US/09015876
Patent No. H001825
GENERAL INFORMATION:
APPLICANT: ROMANO, JOSEPH W
APPLICANT: SHUTLIF, ROYANNE
APPLICANT: WILLIAMS, KIMBERLY G
TITLE OF INVENTION: ISOTHERMAL TRANSCRIPTION BASED ASSAY
TITLE OF INVENTION: FOR THE DETECTION OF HTLV I AND HTLV II
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: AKZO NOBEL PATENT DEPT.
STREET: 1300 PICCARD DRIVE, SUITE 206
CITY: ROCKVILLE
STATE: MARYLAND
COUNTRY: US
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: US/09/015,876
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KLESNER, SHARON N
REGISTRATION NUMBER: 36,335
REFERENCE/DOCKET NUMBER: T/98353
TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-948-7400
TELEFAX: 301-948-9751
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 43 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-015-876-23
Query Match 84.0%; Score 16.8; DB 1; Length 43;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Cy 1 ACTCCTATAGGAGAGATG 20
Db 13 ACTCCTATAGGAGAGATG 32
RESULT 33
US-08-441-887A-24
Sequence 24, Application US/08441887A
Patent No. 5837822
GENERAL INFORMATION:
APPLICANT: Chee, Mark
APPLICANT: Cronin, Maureen T.
APPLICANT: Podor, Stephen P.A.
APPLICANT: Huang, Xiaohua X.
APPLICANT: Rubbell, Earl A.
APPLICANT: Lipshutz, Robert J.
APPLICANT: Lobban, Peter E.
APPLICANT: Morris, MacDonald S.
APPLICANT: Sheldon, Edward M.
TITLE OF INVENTION: Arrays of Nucleic Acid Probes on
NUMBER OF SEQUENCES: 360
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICANT: US/08/441,887A
FILING DATE: 16-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/143,312
FILING DATE: 26-OCT-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/082,937
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Liebeschuetz, Joseph O.
REGISTRATION NUMBER: 37,505
REFERENCE/DOCKET NUMBER: 018547-004160US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-326-2400
TELEFAX: 650-326-2422
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 44 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (probe)
US-08-441-887A-24

Query Match 84.0%; Score 16.8; DB 2; Length 44;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ACTCACTATAGGAGAGATG 20
|||||
DB 8 ACTCACTATAGGAGAGATG 27

RESULT 34
US-08-544-381B-121
; Sequence 121, Application US/08544381B
; Patent No. 6027880
; GENERAL INFORMATION:
; APPLICANT: Cronin, Maureen T.
; APPLICANT: Miyada, Charles Garrett
; APPLICANT: Hubbell, Earl A.
; APPLICANT: Chee, Mark
; APPLICANT: Feder, Stephen P.A.
; APPLICANT: Huang, Xiaohua C.
; APPLICANT: Lipschutz, Robert J.
; APPLICANT: Lobban, Peter S.
; APPLICANT: Morris, Macdonald S.
; APPLICANT: Sheldon, Edward L.
; TITLE OF INVENTION: Arrays of Nucleic Acid Probes for
; TITLE OF INVENTION: Detecting Cystic Fibrosis
; NUMBER OF SEQUENCES: 250
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/544,381B
; FILING DATE: 10-OCT-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/510,521
; FILING DATE: 02-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12305
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/284,064
; FILING DATE: 02-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/143,312
; FILING DATE: 26-OCT-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Liebeschuetz, Joe
; REGISTRATION NUMBER: 37,505
; REFERENCE/DOCKET NUMBER: 018547-004130US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 121:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (primer)
US-08-544-381B-121

Query Match 84.0%; Score 16.8; DB 3; Length 44;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGATG 20
|||||

DB 8 ACTCACTATAGGAGAGATG 27

RESULT 35
US-09-710-200-4
; Sequence 4, Application US/09710200
; Patent No. 6379897
; GENERAL INFORMATION:
; APPLICANT: Nanogen, Inc.
; APPLICANT: Weidenhammer, Elaine M.
; APPLICANT: Wang, Jing
; APPLICANT: Xu, Xiao
; APPLICANT: Heller, Michael J.
; APPLICANT: Kahl, Brenda F.
; TITLE OF INVENTION: IMPROVED METHODS FOR GENE EXPRESSION MONITORING ON ELECTRONIC MI
; FILE REFERENCE: 256/262 Patrick S. Bagleman
; CURRENT APPLICATION NUMBER: US/09/710,200
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotinylated
US-09-710-200-4

Query Match 84.0%; Score 16.8; DB 4; Length 48;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGATG 20
|||||
DB 8 ACTCACTATAGGAGAGATG 27

RESULT 36
US-09-975-408-4
; Sequence 4, Application US/09975408
; Patent No. 6491122
; GENERAL INFORMATION:
; APPLICANT: Nanogen, Inc.
; APPLICANT: Weidenhammer, Elaine M.
; APPLICANT: Xu, Xiao
; APPLICANT: Kahl, Brenda F.
; TITLE OF INVENTION: IMPROVED METHODS FOR GENE EXPRESSION MONITORING ON ELECTRONIC MI
; FILE REFERENCE: 267/174 Patrick S. Bagleman
; CURRENT APPLICATION NUMBER: US/09/975,408
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 09/710,200
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(1)
; OTHER INFORMATION: Biotinylated
US-09-975-408-4

Query Match 84.0%; Score 16.8; DB 4; Length 48;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGATG 20
|||||
DB 8 ACTCACTATAGGAGAGATG 27

```

/ APPLICANT: Nerenberg, Michael I.
/ APPLICANT: Nestin, Lorelei P.
/ APPLICANT: Edman, Carl P.
/ APPLICANT: Carrino, John
/ TITLE OF INVENTION: MULTIPLEX AMPLIFICATION AND SEPARATION OF NUCLEIC ACID
/ TITLE OF INVENTION: SEQUENCES ON A BIOELECTRONIC MICROCHIP USING ASYMMETRIC
/ TITLE OF INVENTION: STRUCTURES
/ FILE REFERENCE: 241/103
/ CURRENT APPLICATION NUMBER: US/09/290,452
/ CURRENT FILING DATE: 1999-04-12
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 16
/ LENGTH: 50
/ TYPE: DNA
/ ORGANISM: human T-cell leukemia virus-1
US-09-290-452-16

Query Match      84.0%; Score 16.8; DB 4; Length 50;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 13 ACTCACTATAGGAGAGATG 32

RESULT 40
US-09-290-338-16
/ Sequence 16, Application US/09290338
/ Patent No. 6326173
/ GENERAL INFORMATION:
/ APPLICANT: Nerenberg, Michael I.
/ APPLICANT: Edman, Carl P.
/ TITLE OF INVENTION: ELECTRONICALLY MEDIATED NUCLEIC ACID
/ TITLE OF INVENTION: AMPLIFICATION IN NASBA
/ FILE REFERENCE: 238/072
/ CURRENT APPLICATION NUMBER: US/09/290,338
/ CURRENT FILING DATE: 1999-04-12
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 16
/ LENGTH: 50
/ TYPE: DNA
/ ORGANISM: human T-cell leukemia virus-1
US-09-290-338-16

Query Match      84.0%; Score 16.8; DB 4; Length 50;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 13 ACTCACTATAGGAGAGATG 32

Search completed: May 24, 2004, 12:31:18
Job time : 31.9434 secs

```

```

RESULT 37
US-09-600-770A-5
/ Sequence 5, Application US/09600770A
/ Patent No. 6489110
/ GENERAL INFORMATION:
/ APPLICANT: Oudshoorn, Pieter
/ APPLICANT: Klatser, Paul
/ TITLE OF INVENTION: BP-TU mRNA AS A MARKER FOR VIABILITY OF BACTERIA
/ FILE REFERENCE: 9250.21
/ CURRENT APPLICATION NUMBER: US/09/600,770A
/ CURRENT FILING DATE: 2000-07-21
/ PRIOR APPLICATION NUMBER: PCT/EP99/00323
/ PRIOR FILING DATE: 1999-01-19
/ NUMBER OF SEQ ID NOS: 95
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 5
/ LENGTH: 49
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURES:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(49)
/ OTHER INFORMATION: Oligonucleotide primer to Escherichia coli BP-TU.
US-09-600-770A-5

Query Match      84.0%; Score 16.8; DB 4; Length 49;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 13 ACTCACTATAGGAGAGATG 32

RESULT 38
US-09-290-577-16
/ Sequence 16, Application US/09290577
/ Patent No. 6238868
/ GENERAL INFORMATION:
/ APPLICANT: Carrino, John J.
/ APPLICANT: Gervu, Louis O.
/ TITLE OF INVENTION: MULTIPLEX AMPLIFICATION AND SEPARATION OF NUCLEIC
/ TITLE OF INVENTION: ACID SEQUENCES USING LIGATION-DEPENDANT STRAND
/ TITLE OF INVENTION: DISPLACEMENT AMPLIFICATION AND BIOELECTRONIC CHIP
/ TITLE OF INVENTION: TECHNOLOGY
/ FILE REFERENCE: 238/238
/ CURRENT APPLICATION NUMBER: US/09/290,577
/ CURRENT FILING DATE: 1999-04-12
/ NUMBER OF SEQ ID NOS: 62
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 16
/ LENGTH: 50
/ TYPE: DNA
/ ORGANISM: human T-cell leukemia virus-1
US-09-290-577-16

Query Match      84.0%; Score 16.8; DB 3; Length 50;
Best Local Similarity 90.0%; Pred. No. 15;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 13 ACTCACTATAGGAGAGATG 32

RESULT 39
US-09-190-452-16
/ Sequence 16, Application US/09290452
/ Patent No. 630823
/ GENERAL INFORMATION:

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 11:41:02 ; Search time 120.755 Seconds
(without alignments)
752.721 Million cell updates/sec

Title: US-10-144-679-2

Perfect score: 20

Sequence: 1 actcactataggagagatg 20

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2953838 seqs, 2272363821 residues

Total number of hits satisfying chosen parameters: 5907676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: 1: Published Applications NA:

2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
4: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq:
6: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq:
8: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
13: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
14: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq:
18: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq:
19: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	16	US-10-144-679-2
2	20	100.0	20	16	US-10-144-679-11
3	20	100.0	20	16	US-10-144-679-81
4	20	100.0	20	17	US-10-144-094-2
5	20	100.0	20	17	US-10-144-094-11
6	20	100.0	20	17	US-10-144-094-81
7	20	100.0	43	16	US-10-144-679-4
8	20	100.0	43	16	US-10-144-679-13
9	20	100.0	43	17	US-10-144-094-4
10	20	100.0	43	17	US-10-144-094-13
11	20	100.0	44	16	US-10-144-679-85
12	20	100.0	97	16	US-10-144-679-6
13	20	100.0	97	16	US-10-144-679-12
14	20	100.0	97	17	US-10-144-094-6

15	20	100.0	97	17	US-10-144-094-12
16	20	100.0	107	16	US-10-144-679-83
17	20	100.0	107	17	US-10-144-094-83
18	19	95.0	19	16	US-10-144-679-79
19	19	95.0	19	17	US-10-144-094-79
20	18.4	92.0	43	16	US-10-144-679-24
21	18.4	92.0	43	17	US-10-144-094-24
22	18.4	92.0	49	13	US-10-461-790-40
23	17.4	87.0	60	9	US-09-870-433-9
24	17.4	87.0	60	9	US-09-893-131-23
25	17.4	87.0	67	9	US-09-870-433-8
26	17.4	87.0	67	9	US-09-893-131-11
27	17.4	87.0	251	17	US-10-257-294-90
28	17.4	87.0	261	17	US-10-257-294-89
29	17.4	87.0	324	17	US-10-257-294-91
30	17.4	87.0	341	17	US-10-257-294-93
31	17	85.0	45	16	US-10-144-679-84
32	17	85.0	45	17	US-10-144-094-84
33	17	85.0	1359	10	US-09-764-891-2309
34	16.8	84.0	44	10	US-09-510-378-121
35	16.8	84.0	48	9	US-09-975-408-4
36	16.8	84.0	48	14	US-10-075-579-4
37	16.8	84.0	48	15	US-10-056-884-72
38	16.8	84.0	48	15	US-10-080-980-73
39	16.8	84.0	48	15	US-10-086-156-56
40	16.8	84.0	48	15	US-10-071-458-33
41	16.8	84.0	48	15	US-10-234-951A-68
42	16.8	84.0	48	15	US-10-005-549-54
43	16.8	84.0	49	10	US-09-963-827B-188
44	16.8	84.0	50	9	US-09-865-807-16
45	16.8	84.0	50	10	US-09-954-594A-16

ALIGNMENTS

RESULT 1

US-10-144-679-2

Sequence 2, Application US/10144679

Publication No. US20030215810A1

GENERAL INFORMATION:

APPLICANT: LIU, JUEYEN

TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON

TITLE OF INVENTION: COLOR CHANGES

FILE REFERENCE: 9800240-0019

CURRENT APPLICATION NUMBER: US/10/144,679

CURRENT FILING DATE: 2002-05-10

NUMBER OF SEQ ID NOS: 88

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 20

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Synthetic chimeric

OTHER INFORMATION: substrate

FEATURE:

OTHER INFORMATION: Description of Combined DNA/RNA Molecule: Synthetic chimeric

OTHER INFORMATION: substrate

US-10-144-679-2

Query Match 100.0%; Score 20; DB 16; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCCTATAGGAGAGATG 20

Db 1 ACTCCTATAGGAGAGATG 20

RESULT 2

US-10-144-679-11

/ Sequence 11, Application US/10144679
/ Publication No. US20030215810A1
/ GENERAL INFORMATION:
/ APPLICANT: LIU, JUEWEN
/ TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
/ FILE REFERENCE: 9800240-0019
/ CURRENT APPLICATION NUMBER: US/10/144,679
/ CURRENT FILING DATE: 2002-05-10
/ NUMBER OF SEQ ID NOS: 88
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 11
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic
/ OTHER INFORMATION: Rh-17DD8
US-10-144-679-11

Query Match 100.0%; Score 20; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 1 ACTCACTATAGGAGAGATG 20

RESULT 3

US-10-144-679-81
/ Sequence 81, Application US/10144679
/ Publication No. US20030215810A1
/ GENERAL INFORMATION:
/ APPLICANT: LIU, JI
/ TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
/ FILE REFERENCE: 9800240-0019
/ CURRENT APPLICATION NUMBER: US/10/144,679
/ CURRENT FILING DATE: 2002-05-10
/ NUMBER OF SEQ ID NOS: 88
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 81
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic chimeric
/ OTHER INFORMATION: substrate
/ OTHER INFORMATION: Description of Combined DNA/RNA Molecule: Synthetic chimeric
US-10-144-679-81

Query Match 100.0%; Score 20; DB 16; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 1 ACTCACTATAGGAGAGATG 20

RESULT 4

US-10-144-094-2
/ Sequence 2, Application US/10144094
/ Publication No. US20040023216A1
/ GENERAL INFORMATION:
/ APPLICANT: LIU, JI
/ TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR

/ FILE REFERENCE: 10322/44
/ CURRENT APPLICATION NUMBER: US/10/144,094
/ CURRENT FILING DATE: 2002-05-10
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Synthetic chimeric
/ OTHER INFORMATION: substrate
/ OTHER INFORMATION: Description of Combined DNA/RNA Molecule: Synthetic chimeric
US-10-144-094-2

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 1 ACTCACTATAGGAGAGATG 20

RESULT 5

US-10-144-094-11
/ Sequence 11, Application US/10144094
/ Publication No. US20040023216A1
/ GENERAL INFORMATION:
/ APPLICANT: LIU, JI
/ TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
/ FILE REFERENCE: 10322/44
/ CURRENT APPLICATION NUMBER: US/10/144,094
/ CURRENT FILING DATE: 2002-05-10
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 11
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Rh-17DD8
US-10-144-094-11

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 1 ACTCACTATAGGAGAGATG 20

RESULT 6

US-10-144-094-81
/ Sequence 81, Application US/10144094
/ Publication No. US20040023216A1
/ GENERAL INFORMATION:
/ APPLICANT: LIU, JI
/ TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
/ FILE REFERENCE: 10322/44
/ CURRENT APPLICATION NUMBER: US/10/144,094
/ CURRENT FILING DATE: 2002-05-10
/ NUMBER OF SEQ ID NOS: 84
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 81
/ LENGTH: 20
/ TYPE: DNA
/ ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic chimeric
FEATURE:
OTHER INFORMATION: substrate
OTHER INFORMATION: Description of Combined DNA/RNA Molecule: Synthetic chimeric
OTHER INFORMATION: substrate
US-10-144-094-81

Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 1 ACTCACTATAGGAGAGATG 20

RESULT 7
US-10-144-679-4
Sequence 4, Application US/10144679
Publication No. US20030215810A1
GENERAL INFORMATION:
APPLICANT: LIU, YI
TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
FILE REFERENCE: 9800240-0019
CURRENT APPLICATION NUMBER: US/10/144, 679
CURRENT FILING DATE: 2002-05-10
NUMBER OF SEQ ID NOS: 88
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 43
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-144-679-4

Query Match 100.0%; Score 20; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 8
US-10-144-679-13
Sequence 13, Application US/10144679
Publication No. US20030215810A1
GENERAL INFORMATION:
APPLICANT: LIU, YI
TITLE OF INVENTION: SIMPLE CATALYTIC DNA BIOSENSORS FOR IONS BASED ON
FILE REFERENCE: 9800240-0019
CURRENT APPLICATION NUMBER: US/10/144, 679
CURRENT FILING DATE: 2002-05-10
NUMBER OF SEQ ID NOS: 88
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 43
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Zn-DNA
US-10-144-679-13

Query Match 100.0%; Score 20; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.9;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 9
US-10-144-094-4
Sequence 4, Application US/10144094
Publication No. US20040023216A1
GENERAL INFORMATION:
APPLICANT: LIU, YI
TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
FILE REFERENCE: 10322/44
CURRENT APPLICATION NUMBER: US/10/144, 094
CURRENT FILING DATE: 2002-05-10
NUMBER OF SEQ ID NOS: 84
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 43
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-144-094-4

Query Match 100.0%; Score 20; DB 17; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 10
US-10-144-094-13
Sequence 13, Application US/10144094
Publication No. US20040023216A1
GENERAL INFORMATION:
APPLICANT: LIU, YI
TITLE OF INVENTION: NEW FLUORESCENCE BASED BIOSENSOR
FILE REFERENCE: 10322/44
CURRENT APPLICATION NUMBER: US/10/144, 094
CURRENT FILING DATE: 2002-05-10
NUMBER OF SEQ ID NOS: 84
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 43
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Zn-DNA
US-10-144-094-13

Query Match 100.0%; Score 20; DB 17; Length 43;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 11
US-10-144-679-85
Sequence 85, Application US/10144679
Publication No. US20030215810A1
GENERAL INFORMATION:
APPLICANT: LIU, YI

US-10-075-579-4
Sequence 4, Application US/10075579
Publication No. US20020119484A1
GENERAL INFORMATION:

; PUBLICATION NO. 0940030036113541
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYPEPTIDE ENCODING A NOVEL HUMAN POTASSIUM CHANNEL BETA-SUBUNIT OF INVENTION: K-POT46. EXPRESSED HIGHLY IN THE SMALL INTESTINE
 ; TITLE OF INVENTION: K-POT46. EXPRESSED HIGHLY IN THE SMALL INTESTINE

FILE REFERENCE: D0121 NP
CURRENT APPLICATION NUMBER: US/10/080,980
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: US 60/270,132
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US 60/278,953
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 74
SOFTWARE: PatentIn version 3.0
SEQ ID NO 73
LENGTH: 48
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-10-080-980-73

Query Match 84.0%; Score 16.8; DB 15; Length 48;
Best Local Similarity 90.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
|||||
Db 12 ACTCACTATAGGAGAGATG 31

RESULT 39
US-10-086-156-56
Sequence 56, Application US/10086156
Publication No. US20030054989A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING TWO NOVEL HUMAN POTASSIUM CHANNEL BETA-SUB
TITLE OF INVENTION: K-betaM4 and K-betaM5
FILE REFERENCE: D0115NP
CURRENT APPLICATION NUMBER: US/10/086,156
CURRENT FILING DATE: 2002-02-28
PRIOR APPLICATION NUMBER: US 60/272,190
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: US 60/274,258
PRIOR FILING DATE: 2001-03-07
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn version 3.0
SEQ ID NO 56
LENGTH: 48
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-10-086-156-56

Query Match 84.0%; Score 16.8; DB 15; Length 48;
Best Local Similarity 90.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
|||||
Db 12 ACTCACTATAGGAGAGATG 31

RESULT 40
US-10-071-458-33
Sequence 33, Application US/10071458
Publication No. US20030114371A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN POTASSIUM CHANNEL BETA-SUB
TITLE OF INVENTION: K-betaM3
FILE REFERENCE: D0114.NP
CURRENT APPLICATION NUMBER: US/10/071,458
CURRENT FILING DATE: 2002-02-07
PRIOR APPLICATION NUMBER: US 60/267,039
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 60/281,224
PRIOR FILING DATE: 2001-04-03
NUMBER OF SEQ ID NOS: 76
SOFTWARE: PatentIn version 3.0

SEQ ID NO 33
LENGTH: 48
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-10-071-458-33

Query Match 84.0%; Score 16.8; DB 15; Length 48;
Best Local Similarity 90.0%; Pred. No. 87;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ACTCACTATAGGAGAGATG 20
|||||
Db 12 ACTCACTATAGGAGAGATG 31

Search completed: May 24, 2004, 13:45:29
Job time : 120.755 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 24, 2004, 10:18:57 ; Search time 589.811 Seconds
(without alignments)
1469.725 Million cell updates/sec
Title: US-10-144-679-2
Perfect score: 20
Sequence: 1 actcactatagaagagatg 20
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ets.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_scs.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_plni.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	6	AX418516	AX418516 Sequence
2	20	100.0	20	6	AX418525	AX418525 Sequence
3	20	100.0	20	6	AX418595	AX418595 Sequence
C 4	20	100.0	20	6	AR038879	AR038879 Sequence
C 5	20	100.0	30	6	AR107371	AR107371 Sequence
C 6	20	100.0	30	6	AR179333	AR179333 Sequence
C 7	20	100.0	30	6	BD191596	BD191596 Enzymatic
8	20	100.0	38	6	AR405843	AR405843 Sequence
9	20	100.0	38	6	AR405888	AR405888 Sequence
10	20	100.0	38	6	AR405892	AR405892 Sequence
11	20	100.0	38	6	BD190590	BD190590 Bioreacti
12	20	100.0	38	6	BD190635	BD190635 Bioreacti
13	20	100.0	38	6	AR038891	AR038891 Sequence
14	20	100.0	43	6	AR038891	AR038891 Sequence
15	20	100.0	43	6	AR107373	AR107373 Sequence
16	20	100.0	43	6	AR107375	AR107375 Sequence
17	20	100.0	43	6	AR179335	AR179335 Sequence
18	20	100.0	43	6	AX418518	AX418518 Sequence
19	20	100.0	43	6	AX418527	AX418527 Sequence
20	20	100.0	43	6	BD191598	BD191598 Enzymatic
21	20	100.0	68	6	AR107385	AR107385 Sequence
22	20	100.0	69	6	AR405868	AR405868 Sequence
23	20	100.0	69	6	BD190615	BD190615 Bioreacti
24	20	100.0	87	6	AR405867	AR405867 Sequence
25	20	100.0	87	6	BD190614	BD190614 Bioreacti
C 26	20	100.0	97	6	AX418520	AX418520 Sequence
27	20	100.0	97	6	AX418526	AX418526 Sequence
28	20	100.0	107	6	AX418597	AX418597 Sequence
29	19	95.0	19	6	AX418593	AX418593 Sequence
30	19	95.0	43	6	AR038893	AR038893 Sequence
31	19	95.0	43	6	AR179337	AR179337 Sequence
32	19	95.0	43	6	BD191600	BD191600 Enzymatic
33	19	95.0	65	6	AR405883	AR405883 Sequence
34	19	95.0	65	6	BD190630	BD190630 Bioreacti
35	19	95.0	107	6	AR038893	AR038893 Sequence
36	19	95.0	107	6	AR179347	AR179347 Sequence
37	19	95.0	107	6	BD191610	BD191610 Enzymatic
38	18.4	92.0	43	6	AX418538	AX418538 Sequence
39	18.4	92.0	48	6	AX777164	AX777164 Sequence
40	18.4	92.0	48	6	AX777194	AX777194 Sequence
41	18.4	92.0	48	6	AX777234	AX777234 Sequence
42	18	90.0	19	6	AR107364	AR107364 Sequence
43	18	90.0	19	6	AR107369	AR107369 Sequence
44	18	90.0	19	6	AR107374	AR107374 Sequence
C 45	18	90.0	103	6	AX080704	AX080704 Sequence

ALIGNMENTS

RESULT 1
AX418516
LOCUS AX418516 20 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 2 from Patent WO0200006.
ACCESSION AX418516
VERSION AX418516.1 GI:21523381
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
1
Lu, Y. and Li, J.
Nucleic acid enzyme biosensor for ions
TITLE Patent: WO 020006-A 2 03-JAN-2002;
JOURNAL THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)

us-10-144-679-2.zgc

Tue May 25 08:19:57 2004

```
FEATURES
  source
    Location/Qualifiers
      1..20
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="Description of Combined DNA/RNA Molecule: Chimeric
      substrate"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 1 ACTCACTATAGGAGAGATG 20

RESULT 2
LOCUS AX418525 20 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 11 from Patent WO2000006.
ACCESSION AX418525
VERSION AX418525.1 GI:21523390
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Lu, Y. and Li, J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 020006-A 11 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
  source
    Location/Qualifiers
      1..20
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="Rh-17DD"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 1 ACTCACTATAGGAGAGATG 20

RESULT 3
LOCUS AX418595 20 bp DNA linear PAT 19-JUN-2002
DEFINITION Sequence 81 from Patent WO020006.
ACCESSION AX418595
VERSION AX418595.1 GI:21523450
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Lu, Y. and Li, J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 020006-A 81 03-JAN-2002;
THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
  source
    Location/Qualifiers
      1..20
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /note="Description of Combined DNA/RNA Molecule: Chimeric
      substrate"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 1 ACTCACTATAGGAGAGATG 20

RESULT 4
LOCUS AR038979 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 9 from patent US 5807718.
ACCESSION AR038979
VERSION AR038979.1 GI:5958342
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Joyce, G.P. and Breaker, R.E.
TITLE Enzymatic DNA molecules
JOURNAL Patent: US 5807718-A 9 15-SEP-1998;
Location/Qualifiers
FEATURES
  source
    Location/Qualifiers
      1..30
      /organism="unknown"
      /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 21 ACTCACTATAGGAGAGATG 2

RESULT 5
LOCUS ARL07371 30 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 9 from patent US 6110462.
ACCESSION ARL07371
VERSION ARL07371.1 GI:12822858
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Barbas, C.F., Joyce, G., Santoro, S.W. and Kandasamy, S.
TITLE Enzymatic DNA molecules that contain modified nucleotides
JOURNAL Patent: US 6110462-A 9 29-AUG-2000;
Location/Qualifiers
FEATURES
  source
    Location/Qualifiers
      1..30
      /organism="unknown"
      /mol_type="unassigned DNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 21 ACTCACTATAGGAGAGATG 2

RESULT 6
LOCUS ARL79333 30 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 9 from patent US 6126174.
ACCESSION ARL79333
```



```

VERSION      ARI79333.1  GI:20220888
KEYWORDS     .
SOURCE       Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 30)
AUTHORS      Joyce, G.P. and Breaker, R.R.
TITLE        Enzymatic DNA molecules.
JOURNAL      Patent: US 6326114-A 9 04-DEC-2001;
FEATURES     Location/Qualifiers
             source
             1..30
             /organism="unknown"
             /mol_type="unassigned DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
   |||||
DB 21 ACTCACTATAGGAGAGATG 2

RESULT 7
LOCUS      BD191596/c      30 bp      DNA      linear      PAT 17-JUL-2003
DEFINITION Enzymatic DNA molecules.
ACCESSION  BD191596
VERSION     BD191596.1  GI:33001335
KEYWORDS   JP 2002514080-A/9.
SOURCE     unidentified
ORGANISM   unclassified.
REFERENCE  1 (bases 1 to 30)
AUTHORS    Joyce, G.P. and Breaker, R.R.
TITLE      Enzymatic DNA molecules
JOURNAL    Patent: JP 2002514080-A 9 14-MAY-2002;
           THE SCRIPPS RESEARCH INSTITUTE
COMMENT    PN JP 2002514080-A/9
           PD 14-MAY-2002
           PP 29-APR-1998 JP 1998547359
           PR 29-APR-1997 US 60/045228
           PT GERALD P JOYCE RONALD R BREAKER
           PC C1201/68 C1209/22 C07H21/04
           CC Strandedness: Single;
           CC Topology: Linear;
           FH Key Location/Qualifiers.
FEATURES    source
             1..30
             /organism="unidentified"
             /mol_type="genomic DNA"
             /db_xref="taxon:32644"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
   |||||
DB 21 ACTCACTATAGGAGAGATG 2

RESULT 8
LOCUS      AR405843      38 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 8 from patent US 6630306.
ACCESSION  AR405843
VERSION     AR405843.1  GI:40154862
KEYWORDS   Unknown.
SOURCE     Unknown.
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
   |||||
DB 9 ACTCACTATAGGAGAGATG 28

RESULT 9
LOCUS      AR405888      38 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 53 from patent US 6630306.
ACCESSION  AR405888
VERSION     AR405888.1  GI:40154907
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 38)
AUTHORS    Breaker, R.R.
TITLE      Bioreactive allosteric polynucleotides
JOURNAL    Patent: US 6630306-A 53 07-OCT-2003;
           Location/Qualifiers
FEATURES     source
             1..38
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
   |||||
DB 14 ACTCACTATAGGAGAGATG 33

RESULT 10
LOCUS      AR405892      38 bp      DNA      linear      PAT 18-DEC-2003
DEFINITION Sequence 57 from patent US 6630306.
ACCESSION  AR405892
VERSION     AR405892.1  GI:40154911
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 38)
AUTHORS    Breaker, R.R.
TITLE      Bioreactive allosteric polynucleotides
JOURNAL    Patent: US 6630306-A 57 07-OCT-2003;
           Location/Qualifiers
FEATURES     source
             1..38
             /organism="unknown"
             /mol_type="genomic DNA"
ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
   |||||
DB 9 ACTCACTATAGGAGAGATG 28

```

```

Qy 1 ACTCACTATAGGAGAGATG 20
Db 14 ACTCACTATAGGAGAGATG 33

RESULT 11
LOCUS BD190590 38 bp DNA PAT 17-JUL-2003
DEFINITION Bioreactive allosteric polynucleotide.
ACCESSION BD190590
VERSION BD190590.1 GI:33000329
KEYWORDS JP 2002514913-A/8.
SOURCE Rattus
ORGANISM Rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
1 (bases 1 to 38)
REFERENCE Braker,R.R.
AUTHORS Bioreactive allosteric polynucleotide.
TITLE Patent: JP 2002514913-A 8 21-MAY-2002;
JOURNAL AILE UNIVERSITY
COMMENT PN JP 2002514913-A/8
PD 21-MAY-2002
PF 18-DEC-1997 JP 1998528049 60/033684,08-AUG-1997 US 60/055039 PI
PR 19-DEC-1996 US 60/033684,08-AUG-1997 US
PC C12N15/09,C12M1/00,C12Q1/68,C12N15/00
RONALD R BRAKER
CC Strandedness: Single;
CC Topology: Linear;
CC Key Location/Qualifiers.
FH Key Location/Qualifiers

FEATURES
source
1..38
/organism="Rattus"
/mol_type="genomic DNA"
/db_xref="taxon:10114"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 14 ACTCACTATAGGAGAGATG 33

RESULT 12
LOCUS BD190635 38 bp DNA PAT 17-JUL-2003
DEFINITION Bioreactive allosteric polynucleotide.
ACCESSION BD190635
VERSION BD190635.1 GI:33000374
KEYWORDS JP 2002514913-A/53.
SOURCE Rattus
ORGANISM Rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
1 (bases 1 to 38)
REFERENCE Braker,R.R.
AUTHORS Bioreactive allosteric polynucleotide.
TITLE Patent: JP 2002514913-A 53 21-MAY-2002;
JOURNAL AILE UNIVERSITY
COMMENT PN JP 2002514913-A/53
PD 21-MAY-2002
PF 18-DEC-1997 JP 1998528049 60/033684,08-AUG-1997 US 60/055039 PI
PR 19-DEC-1996 US 60/033684,08-AUG-1997 US
PC C12N15/09,C12M1/00,C12Q1/68,C12N15/00
RONALD R BRAKER
CC Strandedness: Single;
CC Topology: Linear;
CC Key Location/Qualifiers.
FH Key Location/Qualifiers

FEATURES
source
1..38
/organism="Rattus"
/mol_type="genomic DNA"
/db_xref="taxon:10114"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 14 ACTCACTATAGGAGAGATG 33

RESULT 13
LOCUS BD190639 38 bp DNA PAT 17-JUL-2003
DEFINITION Bioreactive allosteric polynucleotide.
ACCESSION BD190639
VERSION BD190639.1 GI:33000378
KEYWORDS JP 2002514913-A/57.
SOURCE Rattus
ORGANISM Rattus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
1 (bases 1 to 38)
REFERENCE Braker,R.R.
AUTHORS Bioreactive allosteric polynucleotide.
TITLE Patent: JP 2002514913-A 57 21-MAY-2002;
JOURNAL AILE UNIVERSITY
COMMENT PN JP 2002514913-A/57
PD 21-MAY-2002
PF 18-DEC-1997 JP 1998528049 60/033684,08-AUG-1997 US 60/055039 PI
PR 19-DEC-1996 US 60/033684,08-AUG-1997 US
PC C12N15/09,C12M1/00,C12Q1/68,C12N15/00
RONALD R BRAKER
CC Strandedness: Single;
CC Topology: Linear;
CC Key Location/Qualifiers.
FH Key Location/Qualifiers

FEATURES
source
1..38
/organism="Rattus"
/mol_type="genomic DNA"
/db_xref="taxon:10114"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 38;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 14 ACTCACTATAGGAGAGATG 33

RESULT 14
LOCUS AR038981 43 bp DNA PAT 29-SEP-1999
DEFINITION Sequence 11 from patent US 5807718.
ACCESSION AR038981
VERSION AR038981.1 GI:5958344
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 43)
AUTHORS Joyce,G.P. and Breaker,R.R.
TITLE Euzymatic DNA molecules
JOURNAL Patent: US 5807718-A 11 15-SEP-1998;
FEATURES Location/Qualifiers
source
1..43
/organism="unknown"
/mol_type="unassigned DNA"

```

```

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 43;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 15
LOCUS AR107373 43 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 11 from patent US 6110462.
ACCESSION AR107373
VERSION AR107373.1 GI:12822860
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Barbas,C.F., Joyce,G., Santoro,S.W. and Kandasamy,S.
TITLE Enzymatic DNA molecules that contain modified nucleotides
JOURNAL Patent: US 6110462-A 11 29-AUG-2000;
FEATURES
    source
    1..43
    /organism="unknown"
    /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 43;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 16
LOCUS AR107375 43 bp DNA linear PAT 14-FEB-2001
DEFINITION Sequence 13 from patent US 6110462.
ACCESSION AR107375
VERSION AR107375.1 GI:12822862
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Barbas,C.F., Joyce,G., Santoro,S.W. and Kandasamy,S.
TITLE Enzymatic DNA molecules that contain modified nucleotides
JOURNAL Patent: US 6110462-A 13 29-AUG-2000;
FEATURES
    source
    1..43
    /organism="unknown"
    /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 43;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 17
LOCUS AR179335 43 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 11 from patent US 6326174.
ACCESSION AR179335
VERSION AR179335.1 GI:20220890
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Joyce,G.P. and Breaker,R.R.
TITLE Enzymatic DNA molecules
JOURNAL Patent: US 6326174-A 11 04-DEC-2001;
FEATURES
    source
    1..43
    /organism="unknown"
    /mol_type="unassigned DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 43;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 18
LOCUS AX418518 43 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 4 from Patent WO2000006.
ACCESSION AX418518
VERSION AX418518.1 GI:21523383
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Lu,Y. and Li,J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 0200006-A 4 03-JAN-2002;
    THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
    source
    1..43
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:3630"
    /note="Primer"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 43;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 19
LOCUS AX418527 43 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 13 from Patent WO2000006.
ACCESSION AX418527
VERSION AX418527.1 GI:21523392
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Lu,Y. and Li,J.
TITLE Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 0200006-A 13 03-JAN-2002;
    THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES
    source
    1..43
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /note="Primer"

```



```

Query Match      100.0%; Score 20; DB 6; Length 69;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGATG 20
   |||||
DB 14 ACTCACTATAGGAAGAGATG 33

RESULT 24
LOCUS AR405867
DEFINITION Sequence 32 from patent US 6630306.
ACCESSION AR405867
VERSION AR405867.1 GI:40154886
KEYWORDS
SOURCE Unknown.
ORGANISM
REFERENCE
AUTHORS
TITLES
JOURNAL
FEATURES
source
1. .87
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGATG 20
   |||||
DB 14 ACTCACTATAGGAAGAGATG 33

RESULT 25
LOCUS BD190614
DEFINITION Bioreactive allosteric polynucleotide.
ACCESSION BD190614
VERSION BD190614.1 GI:33000353
KEYWORDS JP 2002514913-A/32.
SOURCE Rattus
ORGANISM Rattus
REFERENCE
AUTHORS Braker,R.R.
TITLES Bioreactive allosteric polynucleotide
JOURNAL Patent: JP 2002514913-A 32 21-MAY-2002;
COMMENT PN 2002514913-A/32
PD 21-MAY-2002
PF 18-DEC-1997 JP 1998528049
PR 18-DEC-1996 US 60/033684, 08-AUG-1997 US 60/055039 PI
PC C12N15/08, C12M1/00, C12Q1/68, C12N15/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1. .87
/organism="Rattus"
/mol_type="genomic DNA"
/db_xref="taxon:10114"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 87;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 ACTCACTATAGGAAGAGATG 20
   |||||
DB 14 ACTCACTATAGGAAGAGATG 33

RESULT 26
LOCUS AX418520
DEFINITION Sequence 6 from Patent WO0200006.
ACCESSION AX418520
VERSION AX418520.1 GI:21523385
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
AUTHORS Lu, Y. and Li, J.
TITLES Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 0200006-A 6 03-JAN-2002;
FEATURES THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
source Location/Qualifiers
1. .97
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="DNA Template"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 97;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGATG 20
   |||||
DB 89 ACTCACTATAGGAAGAGATG 70

RESULT 27
LOCUS AX418526
DEFINITION Sequence 12 from Patent WO0200006.
ACCESSION AX418526
VERSION AX418526.1 GI:21523391
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE
AUTHORS Lu, Y. and Li, J.
TITLES Nucleic acid enzyme biosensor for ions
JOURNAL Patent: WO 0200006-A 12 03-JAN-2002;
FEATURES THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
source Location/Qualifiers
1. .97
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Description of Combined DNA/RNA Molecule: Chimeric substrate"

ORIGIN
Query Match      100.0%; Score 20; DB 6; Length 97;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGATG 20
   |||||
DB 9 ACTCACTATAGGAAGAGATG 28

RESULT 28
LOCUS AX418597

```

LOCUS AX418597 107 bp DNA linear PAT 18-JUN-2002
 DEFINITION Sequence 83 from Patent WO0200006.
 ACCESSION AX418597
 VERSION AX418597.1 GI:21523462
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1
 AUTHORS Lu, Y. and Li, J.
 TITLE Nucleic acid enzyme biosensor for ions
 JOURNAL Patent: WO 0200006-A 83 03-JAN-2002;
 THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
 FEATURES
 source 1..107
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /notes="Predicted secondary structure of the G3 deoxyribozyme"

ORIGIN
 Query Match 100.0%; Score 20; DB 6; Length 107;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 C
 QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 Db 19 ACTCACTATAGGAGAGATG 38
 |||||

RESULT 29
 AX418593
 LOCUS 19 bp DNA linear PAT 18-JUN-2002
 DEFINITION Sequence 79 from Patent WO0200006.
 ACCESSION AX418593
 VERSION AX418593.1 GI:21523458
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1
 AUTHORS Lu, Y. and Li, J.
 TITLE Nucleic acid enzyme biosensor for ions
 JOURNAL Patent: WO 0200006-A 79 03-JAN-2002;
 THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
 FEATURES
 source 1..19
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /notes="Description of Combined DNA/RNA Molecule: Chimeric substrate"

ORIGIN
 Query Match 95.0%; Score 19; DB 6; Length 19;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 C
 QY 2 CTCACTATAGGAGAGATG 20
 |||||
 Db 1 CTCACTATAGGAGAGATG 19
 |||||

RESULT 30
 AR038983
 LOCUS 43 bp DNA linear PAT 29-SEP-1999
 DEFINITION Sequence 13 from Patent US 5807718.
 ACCESSION AR038983
 VERSION AR038983.1 GI:5958346
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.

LOCUS BD191500 43 bp DNA linear PAT 17-JUL-2003
 DEFINITION Enzymatic DNA molecules.
 ACCESSION BD191500
 VERSION BD191500.1 GI:33001339
 KEYWORDS JP 200251080-A/13.
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (Bases 1 to 43)
 AUTHORS Joyce, G.F. and Breaker, R.R.
 TITLE Enzymatic DNA molecules
 JOURNAL Patent: JP 200251080-A 13 14-MAY-2002;
 THE SCRIPPS RESEARCH INSTITUTE
 COMMENT
 PN JP 200251080-A/13
 PD 14-MAY-2002
 PF 29-APR-1998 JP 1998547359
 PR 29-APR-1997 US 60/045228
 PI GERALD F JOYCE, RONALD R BREAKER
 PC C1201/48, C12N9/22, C07H21/04
 CC Strandedness: Single;
 CC Topology: Linear;
 CC /Standard name: 'adenosineribonucleotide'

LOCUS BD191500 43 bp DNA linear PAT 17-JUL-2003
 DEFINITION Enzymatic DNA molecules.
 ACCESSION BD191500
 VERSION BD191500.1 GI:33001339
 KEYWORDS JP 200251080-A/13.
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (Bases 1 to 43)
 AUTHORS Joyce, G.F. and Breaker, R.R.
 TITLE Enzymatic DNA molecules
 JOURNAL Patent: JP 200251080-A 13 14-MAY-2002;
 THE SCRIPPS RESEARCH INSTITUTE
 COMMENT
 PN JP 200251080-A/13
 PD 14-MAY-2002
 PF 29-APR-1998 JP 1998547359
 PR 29-APR-1997 US 60/045228
 PI GERALD F JOYCE, RONALD R BREAKER
 PC C1201/48, C12N9/22, C07H21/04
 CC Strandedness: Single;
 CC Topology: Linear;
 CC /Standard name: 'adenosineribonucleotide'

Unclassified.
 1 (Bases 1 to 43)
 Joyce, G.F. and Breaker, R.R.
 TITLE Enzymatic DNA molecules
 JOURNAL Patent: US 5807718-A 13 15-SEP-1998;
 FEATURES
 source 1..43
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 95.0%; Score 19; DB 6; Length 43;
 Best Local Similarity 95.0%; Pred. No. 69;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 C
 QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 Db 19 ACTCACTATAGGAGAGATG 38
 |||||

RESULT 31
 AR179337
 LOCUS 43 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 13 from patent US 6326174.
 ACCESSION AR179337
 VERSION AR179337.1 GI:20220892
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (Bases 1 to 43)
 AUTHORS Joyce, G.F. and Breaker, R.R.
 TITLE Enzymatic DNA molecules
 JOURNAL Patent: US 6326174-A 13 04-DEC-2001;
 FEATURES
 source 1..43
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 95.0%; Score 19; DB 6; Length 43;
 Best Local Similarity 95.0%; Pred. No. 69;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 C
 QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 Db 19 ACTCACTATAGGAGAGATG 38
 |||||

LOCUS BD191500 43 bp DNA linear PAT 17-JUL-2003
 DEFINITION Enzymatic DNA molecules.
 ACCESSION BD191500
 VERSION BD191500.1 GI:33001339
 KEYWORDS JP 200251080-A/13.
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (Bases 1 to 43)
 AUTHORS Joyce, G.F. and Breaker, R.R.
 TITLE Enzymatic DNA molecules
 JOURNAL Patent: JP 200251080-A 13 14-MAY-2002;
 THE SCRIPPS RESEARCH INSTITUTE
 COMMENT
 PN JP 200251080-A/13
 PD 14-MAY-2002
 PF 29-APR-1998 JP 1998547359
 PR 29-APR-1997 US 60/045228
 PI GERALD F JOYCE, RONALD R BREAKER
 PC C1201/48, C12N9/22, C07H21/04
 CC Strandedness: Single;
 CC Topology: Linear;
 CC /Standard name: 'adenosineribonucleotide'

LOCUS BD191500 43 bp DNA linear PAT 17-JUL-2003
 DEFINITION Enzymatic DNA molecules.
 ACCESSION BD191500
 VERSION BD191500.1 GI:33001339
 KEYWORDS JP 200251080-A/13.
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (Bases 1 to 43)
 AUTHORS Joyce, G.F. and Breaker, R.R.
 TITLE Enzymatic DNA molecules
 JOURNAL Patent: JP 200251080-A 13 14-MAY-2002;
 THE SCRIPPS RESEARCH INSTITUTE
 COMMENT
 PN JP 200251080-A/13
 PD 14-MAY-2002
 PF 29-APR-1998 JP 1998547359
 PR 29-APR-1997 US 60/045228
 PI GERALD F JOYCE, RONALD R BREAKER
 PC C1201/48, C12N9/22, C07H21/04
 CC Strandedness: Single;
 CC Topology: Linear;
 CC /Standard name: 'adenosineribonucleotide'

```

FEATURES
  source
    FH Key Location/Qualifiers.
    1..43 Location/Qualifiers
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"
  ORIGIN
    Query Match 95.0%; Score 19; DB 6; Length 43;
    Best Local Similarity 95.0%; Pred. No. 69;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 1 ACTCACTATAGGAGAGATG 20
    |||||
  DB 19 ACTCACTATAGGAGAGATG 38
    |||||

RESULT 33
AR405883
LOCUS AR405883 65 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 48 from patent US 6630306.
ACCESSION AR405883
VERSION AR405883.1 GI:40154902
KEYWORDS
SOURCE Unknown.
ORGANISM
  Unclassified.
  1 (bases 1 to 65)
  AUTHOR Braker, R.R.
  TITLE Bioreactive allosteric polynucleotides
  JOURNAL Patent: US 6630306-A 48 07-OCT-2003;
  FEATURES
    source
      1..65
      /organism="unknown"
      /mol_type="genomic DNA"
  ORIGIN
    Query Match 95.0%; Score 19; DB 6; Length 65;
    Best Local Similarity 95.0%; Pred. No. 66;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 1 ACTCACTATAGGAGAGATG 20
    |||||
  DB 3 ACTCACTATAGGAGAGATG 22
    |||||

RESULT 34
BD190630
LOCUS BD190630 65 bp DNA linear PAT 17-JUL-2003
DEFINITION Bioreactive allosteric polynucleotide.
ACCESSION BD190630
VERSION BD190630.1 GI:33000369
KEYWORDS
SOURCE Rattus
ORGANISM
  Rattus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae.
  1 (bases 1 to 65)
  AUTHOR Braker, R.R.
  TITLE Bioreactive allosteric polynucleotide
  JOURNAL Patent: JP 2002514913-A 48 21-MAY-2002;
  COMMENT
    EN JP 2002514913-A/48
    PD 21-MAY-2002
    PF 18-DEC-1997 JP 1998528049
    PR 19-DEC-1996 US 60/033684, 08-AUG-1997 US 60/055039 PI
    RONALD R BRAKER
    PC C12N15/09, C12M1/00, C12Q1/69, C12N15/00
    CC Strandedness: Single;
    CC Topology: Linear;
    CC N is an RNA A linkage
  PH Key Location/Qualifiers.
  FEATURES
    source
      1..65
      /organism="unidentified"
      /mol_type="genomic DNA"
  ORIGIN
    Query Match 95.0%; Score 19; DB 6; Length 107;
    Best Local Similarity 95.0%; Pred. No. 62;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 1 ACTCACTATAGGAGAGATG 20
    |||||
  DB 19 ACTCACTATAGGAGAGATG 38
    |||||

RESULT 35
AR038993
LOCUS AR038993 107 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 23 from patent US 5807718.
ACCESSION AR038993
VERSION AR038993.1 GI:5958356
KEYWORDS
SOURCE Unknown.
ORGANISM
  Unclassified.
  1 (bases 1 to 107)
  AUTHOR Joyce, G.P. and Breaker, R.R.
  TITLE Enzymatic DNA molecules
  JOURNAL Patent: US 5807718-A 23 15-SEP-1998;
  FEATURES
    source
      1..107
      /organism="unknown"
      /mol_type="unassigned DNA"
  ORIGIN
    Query Match 95.0%; Score 19; DB 6; Length 107;
    Best Local Similarity 95.0%; Pred. No. 62;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 1 ACTCACTATAGGAGAGATG 20
    |||||
  DB 19 ACTCACTATAGGAGAGATG 38
    |||||

RESULT 36
AR179347
LOCUS AR179347 107 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 23 from patent US 6326174.
ACCESSION AR179347
VERSION AR179347.1 GI:20220902
KEYWORDS
SOURCE Unknown.
ORGANISM
  Unclassified.
  1 (bases 1 to 107)
  AUTHOR Joyce, G.P. and Breaker, R.R.
  TITLE Enzymatic DNA molecules
  JOURNAL Patent: US 6326174-A 23 04-DEC-2001;
  FEATURES
    source
      1..107
      /organism="unknown"
      /mol_type="unassigned DNA"
  ORIGIN
    Query Match 95.0%; Score 19; DB 6; Length 107;
    Best Local Similarity 95.0%; Pred. No. 62;
    Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
  QY 1 ACTCACTATAGGAGAGATG 20
    |||||
  DB 19 ACTCACTATAGGAGAGATG 38
    |||||

```



```

RESULT 37
BD191610
LOCUS      BD191610      107 bp      DNA      linear      EAT 17-JUL-2003
DEFINITION Enzymatic DNA molecules.
ACCESSION  BD191610
VERSION    BD191610.1 GI:33001349
KEYWORDS   JP 2002514080-A/23.
SOURCE     unidentified
ORGANISM   unidentified
REFERENCE  1 (bases 1 to 107)
AUTHORS   Joyce,G.F. and Breaker,R.R.
TITLE     Enzymatic DNA molecules
JOURNAL   Patent: JP 2002514080-A 23 14-MAY-2002;
          THE SCRIPPS RESEARCH INSTITUTE
COMMENT   PN JP 2002514080-A/23
          PD 14-MAY-2002
          PF 29-APR-1998 JP 1998547359
          PR 29-APR-1997 US 60/045228
          PI GERALD F JOYCE, RONALD R BREAKER
          PC C19Q1/68, C12H9/22, C07H21/04
          CC Strandedness: Single;
          CC Topology: Linear;
          CC /Standard name: 'adenosineribonucleotide'
          CC /Label = rA
          FH Key Location/Qualifiers.
FEATURES   source
            1..107
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32630"
ORIGIN
Query Match      95.0%; Score 19; DB 6; Length 107;
Best Local Similarity 95.0%; Pred. NO. 62;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACTCACTATAGGAGAGATG 38

RESULT 38
AX418538
LOCUS      AX418538      43 bp      DNA      linear      PAT 18-JUN-2002
DEFINITION Sequence 24 from Patent WO200006.
ACCESSION  AX418538
VERSION    AX418538.1 GI:21523403
KEYWORDS   synthetic construct
SOURCE     synthetic construct
ORGANISM   synthetic construct
REFERENCE  1
AUTHORS   Lu, Y. and Li, J.
TITLE     Nucleic acid enzyme biosensor for ions
JOURNAL   Patent: WO 020006-A 24 03-JAN-2002;
          THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES   source
            1..43
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Co-DNA"
ORIGIN
Query Match      92.0%; Score 18.4; DB 6; Length 43;
Best Local Similarity 95.0%; Pred. NO. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 19 ACCCACTATAGGAGAGATG 38

```

```

RESULT 39
AX777164
LOCUS      AX777164      48 bp      DNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 18 from Patent WO03040301.
ACCESSION  AX777164
VERSION    AX777164.1 GI:32694312
KEYWORDS   synthetic construct
SOURCE     synthetic construct
ORGANISM   synthetic construct
REFERENCE  1
AUTHORS   Deak, P., Frenz, L., Glover, D. and Midgley, C.
TITLE     Cell cycle progression proteins
JOURNAL   Patent: WO 03040301-A 18 15-MAY-2003;
          Cyclacel Limited (GB)
FEATURES   source
            1..48
            Location/Qualifiers
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Primer"
ORIGIN
Query Match      92.0%; Score 18.4; DB 6; Length 48;
Best Local Similarity 95.0%; Pred. NO. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 8 ACTCACTATAGGAGAGATG 27

RESULT 40
AX777194
LOCUS      AX777194      48 bp      DNA      linear      PAT 14-JUL-2003
DEFINITION Sequence 48 from Patent WO03040301.
ACCESSION  AX777194
VERSION    AX777194.1 GI:32694342
KEYWORDS   synthetic construct
SOURCE     synthetic construct
ORGANISM   synthetic construct
REFERENCE  1
AUTHORS   Deak, P., Frenz, L., Glover, D. and Midgley, C.
TITLE     Cell cycle progression proteins
JOURNAL   Patent: WO 03040301-A 48 15-MAY-2003;
          Cyclacel Limited (GB)
FEATURES   source
            1..48
            Location/Qualifiers
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Primer"
ORIGIN
Query Match      92.0%; Score 18.4; DB 6; Length 48;
Best Local Similarity 95.0%; Pred. NO. 1.4e+02;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
DB 8 ACTCACTATAGGAGAGATG 27

```

```

Search completed: May 24, 2004, 11:40:57
Job time : 591.311 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 07:43:51 ; Search time 123.019 seconds
(without alignments)
690.658 Million cell updates/sec

Title: US-10-144-679-2

Perfect score: 20

Sequence: 1 actcactataggagagatg 20

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: N_Geneseq_29Jan04.*

1: Geneseq1980s.*

2: Geneseq1990s.*

3: Geneseq2000s.*

4: Geneseq2001as.*

5: Geneseq2001bs.*

6: Geneseq2002s.*

7: Geneseq2003as.*

8: Geneseq2003bs.*

9: Geneseq2003cs.*

10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	6	AAL45381
2	20	100.0	20	6	AAL45460
3	20	100.0	30	2	AAV82930
4	20	100.0	30	3	AAV82930
5	20	100.0	33	6	AAV82230
6	20	100.0	33	6	AAV82230
7	20	100.0	33	6	AAV82230
8	20	100.0	38	2	AAV36574
9	20	100.0	38	2	AAV36574
10	20	100.0	38	2	AAV36574
11	20	100.0	38	2	AAV36574
12	20	100.0	43	2	AAV82934
13	20	100.0	43	2	AAV82934
14	20	100.0	43	3	AAV82232
15	20	100.0	43	3	AAV82234
16	20	100.0	43	6	AAV45383
17	20	100.0	43	6	AAV45391
18	20	100.0	69	2	AAV36554
19	20	100.0	87	2	AAV36553
20	20	100.0	97	6	AAV45390
21	20	100.0	97	6	AAV45390
22	20	100.0	107	2	AAV34901
23	20	100.0	107	2	AAV82944

24	20	100.0	107	3	AAV82244	Synthetic
25	20	100.0	107	6	AAV45461	G3 deoxyr
26	20	100.0	123	6	AAV10821	Self-cleav
27	20	100.0	123	6	AAV10822	Self-cleav
28	19	95.0	19	6	AAV45458	RNA-cleav
29	19	95.0	65	2	AAV36569	Self-cleav
30	18.4	92.0	43	6	AAV45402	Cobalt-de
31	18	90.0	19	2	AAV82933	Enzymatic
32	18	90.0	19	3	AAV82228	Primer Ol
33	18	90.0	19	3	AAV82223	Terminati
34	18	90.0	19	3	AAV82233	Substrate
35	18	90.0	103	4	AAV60641	Probe #13
36	18	90.0	109	5	AAV60585	Probe #3
37	18	90.0	111	5	AAV60590	Probe #5
38	18	90.0	111	5	AAV60604	Probe #15
39	18	90.0	111	5	AAV60596	Probe #9
40	18	90.0	111	5	AAV60594	Probe #8
41	17.4	87.0	41	4	AAV88919	Escherich
42	17.4	87.0	54	2	AAV28399	Primer fo
43	17.4	87.0	60	4	AAV76370	T7 promot
44	17.4	87.0	60	6	AAV27673	Promoter
45	17.4	87.0	67	4	AAV76369	T7 promot

ALIGNMENTS

RESULT 1						
AAV45381						
ID	AAV45381	standard; DNA; 20 BP.				
AC	AAV45381					
XX	AAV45381					
DT	06-JUN-2002	(first entry)				
DE	Zn(II)-dependent trans-cleaving deoxyribozyme 17E.					
KE	Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;					
KW	quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.					
XX	Unidentified.					
OS	Unidentified.					
XX	Key	Location/Qualifiers				
FT	misc_binding	1..3				
FT		/tag= a				
FT		/bound_moiety= "ribozyme 17E"				
FT		/note= "binds nucleotides 32-34 of ribozyme 17E shown in AAV45380"				
FT	misc_RNA	10				
FT		/tag= b				
FT	misc_binding	12..20				
FT		/tag= c				
FT		/bound_moiety= "ribozyme 17E"				
FT		/note= "binds nucleotides 1-9 of ribozyme 17B shown in AAV45380"				
XX	WO2002000006-A2.					
XX	03-JAN-2002.					
XX	27-JUN-2001; 2001WO-US020557.					
XX	27-JUN-2000; 2000US-03605558.					
XX	(UNII) UNIV ILLINOIS FOUND.					
XX	Lu Y, Li J;					
XX	WPI; 2002-130823/17.					
XX	New nucleic acid enzyme biosensor, useful for the sensitive and					
XX	selective detection of ions, particularly metal ions e.g. lead ions, and					
XX	for determining the concentration of a particular ion in a solution.					

XX PS Claim 18; Fig 5; 57pp; English.

CC The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependent on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a ribozyme
 CC which may be used as a biosensor of the invention

XX SQ Sequence 20 BP; 8 A; 3 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.3;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTCACTATAGGAGAGATG 20
 DB 1 ACTCACTATAGGAGAGATG 20

RESULT 2

AAAL4546C
 ID AAL45460 standard; DNA; 20 BP.
 AC AAL45460;
 DT 06-JUN-2002 (first entry)
 DE RNA-cleaving deoxyribozyme substrate #6.
 KW Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
 KW quencher; fluorophore; photodetector; ion concentration; ribozyme;
 KW substrate; ss.
 OS Unidentified.

Key Location/Qualifiers
 PH misc_binding 1..9 /tag= a
 PT /bound moiety= "RNA-cleaving deoxyribozyme"
 PT /note= "binds nucleotides 33-25 of the deoxyribozyme
 PT shown in AAL45459"
 FT misc_RNA 10 /tag= b
 FT misc_binding 12..20 /tag= c
 FT /bound moiety= "RNA-cleaving deoxyribozyme"
 FT /note= "binds nucleotides 9-1 of the deoxyribozyme shown
 FT in AAL45459"

XX PN WO200300006-A2.

XX PD 03-JAN-2002.

XX PP 27-JUN-2001; 2001WO-US020557.

XX PR 27-JUN-2000; 2000US-00605558.

XX PA (UNII) UNIV ILLINOIS FOUND.

XX PI Lu Y, Li J;

XX DR WPI; 2002-130823/17.

XX FT New nucleic acid enzyme biosensors, useful for the sensitive and
 FT selective detection of ions, particularly metal ions e.g. lead ions, and
 FT for determining the concentration of a particular ion in a solution.

XX PS Example 1; Fig 6; 57pp; English.

CC The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependent on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a substrate of
 CC a ribozyme which may be used as a biosensor of the invention

XX SQ Sequence 20 BP; 8 A; 3 C; 5 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.3;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ACTCACTATAGGAGAGATG 20
 DB 1 ACTCACTATAGGAGAGATG 20

RESULT 3

AAV82930/c
 ID AAV82930 standard; DNA; 30 BP.
 AC AAV82930;
 DT 05-MAR-1999 (first entry)
 DE Enzymatic DNA 30mer template.
 KW Enzyme; catalysis; cleavage; target; pharmaceutical; medical; substrate;
 KW regulator; detergent; dental hygiene; meat tenderiser; ss.
 OS Synthetic.
 XX WO9849346-A1.
 XX 05-NOV-1998.
 XX 29-APR-1998; 98WO-US008677.
 XX 29-APR-1997; 97US-0045228P.
 XX (SCRI) SCRIPPS RES INST.
 XX Joyce GP, Breaker RR;
 XX WPI; 1999-034670/03.

XX PT New catalytic DNA molecules - having site-specific endonuclease activity
 XX in a substrate nucleic acid, used for cleaving target nucleic acid
 XX sequences.

XX PS Disclosure; Page 62; 161pp; English.

XX CC This sequence is used in a method which involves the production of
 XX catalytic DNA molecules which can be used for cleaving target nucleic
 XX acid molecules. Such DNA molecules can be used in pharmaceutical and
 XX medical products (e.g. for wound debridement, clot dissolution), as well
 XX as in household items (e.g. detergents, dental hygiene products, meat
 XX tenderisers). Other suitable substrates include those comprising or
 XX produced by picornaviruses, hepatitisviridae (e.g. HAV, HCV),
 XX papillomaviruses (e.g. HPV), gammaherpesviridae (e.g. HTLV-1 and 11),
 XX lymphocryptoviruses, leukemia viruses (e.g. HTLV-1 and 11),
 XX flaviviruses, togaviruses, herpesviruses (including alphaherpesviruses and
 XX betaherpesviruses), cytomegaloviruses (CMV), influenza viruses, viruses
 XX and retroviruses contributing to immunodeficiency diseases and syndromes
 XX (e.g. HIV-1 and -2), simian and feline immunodeficiency viruses and
 XX bovine leukemia viruses. They can also be used as regulators of gene
 XX expression

XX SQ Sequence 30 BP; 5 A; 10 C; 6 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 DB 21 ACTCACTATAGGAGAGATG 2

RESULT 4
 ID AAA92230 standard; DNA; 30 BP.
 XX
 AC AAA92230;
 XX
 DT 09-JAN-2001 (first entry)
 XX
 DE Template oligonucleotide sequence SEQ ID NO:9.
 XX
 KW Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;
 KW substrate; template; primer; cleavage; hydrolytic cleavage;
 KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.
 XX
 OS Synthetic.
 XX
 PN US6110462-A.
 XX
 PD 29-AUG-2000.
 XX
 PP 03-MAR-1999; 99US-00262142.
 XX
 PR 03-MAR-1999; 99US-00262142.
 XX
 PA (SCKI) SCRIPPS RES INST.
 XX
 PI Barbas CP, Kandasamy S, Joyce G, Santoro SM;
 XX
 DR WPI; 2000-593449/56.
 XX
 XX Enzymatic DNA molecules containing modified nucleotides useful for
 PT cleavage of RNA at a specified position.
 XX
 PS Example 2; Col 29; 52pp; English.
 CC The present invention describes a catalytic DNA molecule that
 CC specifically cleaves a substrate nucleic acid at a defined cleavage site,
 CC where the catalytic DNA molecule comprises at least one pyrimidine
 CC nucleotide (T). (I) is a catalytic DNA molecule, capable of hydrolytic
 CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule
 CC is used to cleave RNA of almost any sequence. This is useful for cleavage
 CC of single-stranded nucleic acid in the absence of a restriction
 CC endonuclease site at a specific position. This is useful in methods of
 CC e.g. cloning and genetic engineering. The catalytic core of the minimised
 CC enzyme is composed of only 12 residues, making this one of the smallest
 CC nucleic acid catalysts known. The catalytic core forms a compact hairpin
 CC structure displaying the 3 imidazole-containing residues. The enzyme can
 CC be made to cleave RNA of almost any sequence by simple alteration of the
 CC two substrate-recognition domains that surround the catalytic core. The
 CC enzyme operates with multiple turnover in the presence of micromolar
 CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic
 CC rate enhancement of approximately 1000000-fold compared to the
 CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the
 CC substrate-recognition properties of nucleic acid enzymes and the chemical
 CC functionality of protein enzymes in a molecule that is small in size, yet
 CC versatile and catalytically efficient. The present sequence represents an
 CC oligonucleotide sequence which is used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 30 BP; 5 A; 10 C; 6 G; 9 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 3; Length 30;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 DB 25 ACTCACTATAGGAGAGATG 6

QY 1 ACTCACTATAGGAGAGATG 20
 DB 21 ACTCACTATAGGAGAGATG 2

RESULT 5
 ID ABK10823/c
 XX
 AC ABK10823 standard; DNA; 33 BP.
 XX
 AC ABK10823;
 XX
 DT 21-MAY-2002 (first entry)
 XX
 DE Template DNA for self-capping deoxyribozyme selection.
 XX
 KW Self-capping deoxyribozyme; conjoined polynucleotide; catalyst; ribozyme;
 KW ss.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT misc_binding 1..33
 FT /tag= a
 FT /bound moiety= "Self-phosphorylating deoxyribozyme"
 FT /note= "Forms a double stranded region with bases 38-6 of
 FT the sequence in ABK10821"
 XX
 PN W0200129249-A2.
 XX
 PD 26-APR-2001.
 XX
 PP 13-OCT-2000; 2000WO-US028508.
 XX
 PR 15-OCT-1999; 99US-0159808P.
 XX
 PA (UYVA) UNIV YALE.
 XX
 PI Breaker RR;
 XX
 DR WPI; 2002-226120/28.
 XX
 XX New conjoined polynucleotides comprising a kinase domain and an adenylase
 PT domain, useful in DNA cloning, for carrying out sequential polynucleotide
 PT manipulations in a serial fashion, or for polynucleotide manipulations.
 PS Example 1; Fig 5B; 55pp; English.
 CC The invention describes a conjoined polynucleotide comprising at least
 CC two catalytic domains which function in concert to provide a chemical
 CC transformation involving multiple sequential reactions. The use of the
 CC conjoined polynucleotides as catalysts offer advantages over protein-
 CC based enzymes in cloning and in a number of commercial and industrial
 CC processes. Conjoined polynucleotides made from DNA are expected to be
 CC much more stable, and can be easily made by automated oligonucleotide
 CC synthesis, and DNA is significantly more resistant to hydrolytic
 CC degradation compared to RNA. Conjoined DNA and RNA may be selected for
 CC their ability to function on solid support and are expected to retain
 CC their activity when immobilised. This sequence represents a template
 CC strand used in the selection protocol for self-capping deoxyribozymes,
 CC described in the invention
 XX
 SQ Sequence 33 BP; 6 A; 8 C; 7 G; 12 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 DB 25 ACTCACTATAGGAGAGATG 6

PR 19-DEC-1996; 96US-0033694P.
 PR 08-AUG-1997; 97US-0055039P.
 PA (UYVA) UNIV YALE.
 PI Breaker RR;
 XX WPI, 1998-362715/31.
 DR Allosteric polynucleotides - having modified function or configuration,
 PT used for the production of biosensors.
 XX Example 3; Page 47; 106pp; English.
 CC PCR primers AAV36576-78 were used to amplify the DNA pool, and selected
 CC self-cleaving DNAs (deoxyribozymes). The specification describes an
 CC allosteric DNA polynucleotide which can modify a function or
 CC configuration of the polynucleotide with a chemical effector and/or a
 CC physical signal. The allosteric polynucleotides can be used for detecting
 CC the presence or concentrations of compounds such as amino acids,
 CC peptides, nucleosides, nucleotides, steroids, steroids, microbial or cellular
 CC metabolites, blood or urine components, pharmaceuticals, pesticides,
 CC herbicides, or food toxins. The allosteric polynucleotides can also be
 CC used for detecting physical signals such as radiation and temperature
 CC changes. The polynucleotides can also be used in biosensors, in which
 CC they are more stable than proteins
 XX Sequence 38 BP; 14 A; 7 C; 9 G; 8 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 20; DB 2; Length 38;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ACTCATTATAGGAGAGATG 20
 DB 14 ACTCATTATAGGAGAGATG 33
 RESULT 9
 AAV36529
 ID AAV36529 standard; DNA; 38 BP.
 AC AAV36529;
 XX 09-OCT-1998 (first entry)
 DT PCR primer 2 used in the course of the invention.
 DE Allosteric DNA polynucleotide; modify; function; configuration;
 KW detection; biosensor; PCR primer; ss.
 XX Synthetic.
 OS WO9827104-A1.
 PN 25-JUN-1998.
 PD 18-DEC-1997; 97WO-US024158.
 PF 19-DEC-1996; 96US-0033694P.
 PR 08-AUG-1997; 97US-0055039P.
 XX (UYVA) UNIV YALE.
 PA Breaker RR;
 PI WPI, 1998-362715/31.
 DR Allosteric polynucleotides - having modified function or configuration,
 PT used for the production of biosensors.
 XX Example 2; Fig 9B; 106pp; English.

CC PCR primers AAV36528-29 were used in the course of the invention. The
 CC specification describes an allosteric DNA polynucleotide which can modify
 CC a function or configuration of the polynucleotide with a chemical
 CC effector and/or a physical signal. The allosteric polynucleotides can be
 CC used for detecting the presence or concentrations of compounds such as
 CC amino acids, peptides, nucleosides, nucleotides, steroids, microbial or
 CC cellular metabolites, blood or urine components, pharmaceuticals,
 CC pesticides, herbicides, or food toxins. The allosteric polynucleotides
 CC can also be used for detecting physical signals such as radiation and
 CC temperature changes. The polynucleotides can also be used in biosensors,
 CC in which they are more stable than proteins
 XX Sequence 38 BP; 14 A; 7 C; 9 G; 8 T; 0 U; 0 Other;
 SQ Query Match 100.0%; Score 20; DB 2; Length 38;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ACTCATTATAGGAGAGATG 20
 DB 14 ACTCATTATAGGAGAGATG 33
 RESULT 10
 AAV69804
 ID AAV69804 standard; DNA; 38 BP.
 AC AAV69804;
 XX 01-FEB-1999 (first entry)
 DT PCR extension reaction primer 2 used in in vitro selection.
 DE Hammerhead ribozyme; nucleic acid catalyst; synthesis; target;
 KW ATP-dependent allosteric ribozyme; endonuclease activity; cleave;
 KW modulating gene expression; treatment; diagnosis; genetic drift;
 KW mutation; cell proliferation; PCR primer; ss.
 XX Synthetic.
 OS WO9843993-A2.
 PN 08-OCT-1998.
 PD 30-MAR-1998; 98WO-US006231.
 PF 31-MAR-1997; 97US-0042905P.
 PR (UYVA) UNIV YALE.
 PA Breaker RR;
 PI WPI, 1998-568274/48.
 DR New nucleic acid with endonuclease activity - used to cleave targetted
 XX RNA for modulating gene expression in plant or animal cells, particularly
 XX for treatment or diagnosis of disease.
 XX Example 1; Page 37; 61pp; English.
 CC The present invention describes nucleic acids having endonuclease
 CC activity. The nucleic acids of the invention can cleave a target nucleic
 CC acid, particularly RNA, for modulation of gene expression in plant or
 CC animal cells, e.g. for diagnosis and/or treatment of diseases, such as
 CC inhibition of cell proliferation, and also for examining genetic drift
 CC and mutations in cells, to detect the target (mutant or wild-type) in
 CC cells and possibly for RNA sequencing. The nucleic acids, optionally
 CC complexed with cationic lipid, are delivered to smooth muscle cells (via
 CC catheter or stent, incorporated in biopolymer and by injection), or
 CC vectors that express them are introduced into cells ex vivo or in vivo.
 CC The nucleic acids catalyze both intra- and inter-molecular endonuclease
 CC reactions in a sequence-specific manner, and are not homologous with
 CC known ribozymes. They can be designed to target almost any RNA transcript

CC and since they are relatively small are reasonably inexpensive to
 CC produce. The present sequence represents a PCR extension reaction primer
 CC used in an example from the present invention for in vitro selection

XX Sequence 38 BP; 14 A; 7 C; 9 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 38;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 DB 14 ACTCACTATAGGAGAGATG 33

RESULT 11

AA134973
 ID AAT34973 standard; DNA; 43 BP.

XX AC AAT34973;

XX DT 11-FEB-1997 (first entry)

XX DE Substrate for ssDNA molecules having site-specific cleavage activity.

XX EN Endonuclease; cleavage; RNA; catalyst; wound debridement;

XX KW Clot dissolution; meat tenderizer; viral infection; dental hygiene;

XX KW detergent; enzyme isolation; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT misc_feature 28

XX FT /*tag= a
 /*note= "adenosine ribonucleotide; cleavage site"

XX DN WC9617086-A1.

XX PD 06-JUN-1996.

XX PP 01-DEC-1995; 95WO-US015580.

XX PR 02-DEC-1994; 94US-00349023.

XX PR 07-JUN-1995; 95US-00472194.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Joyce GF, Breaker RR;

XX DR WPI; 1996-286834/29.

XX PT New DNA with site specific endonuclease activity - partic. for cleavage
 of RNA, e.g. for medical use.

XX PS Example 2; Page 62; 114pp; English.

XX CC AAT34971 is a substrate for single stranded (ss) DNA molecules (see
 CC AAT34892-900) that have site-specific endonuclease activity. Such
 CC catalytic DNA molecules (see also AAT34901-T34967) specifically cleave ss
 CC nucleic acids, esp. RNA sequences (this sequence is cleaved at the of the
 CC ribonucleotide base) site. The catalytic DNAs may be used in medicine
 CC e.g. for wound debridement, clot dissolution, etc., or in detergents,
 CC dental hygiene products and meat tenderizers. The DNAs may be useful to
 CC treat viral infections such as HIV when targeted to viral nucleic acid
 CC and they may be expressed in a target host cell. In vitro selection of
 CC the DNAs allows the isolation of catalysts/enzymes without the need for
 CC prior knowledge of their compen. or structure

XX Sequence 43 BP; 15 A; 8 C; 12 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 43;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 DB 19 ACTCACTATAGGAGAGATG 38

RESULT 12

AAV82934
 ID AAV82934 standard; DNA; 43 BP.

XX AC AAV82934;

XX DT 05-MAR-1999 (first entry)

XX DE Enzymatic DNA fixed substrate domain.

XX EN Enzyme; catalysis; cleavage; target; pharmaceutical; medical; substrate;

XX KW regulator; detergent; dental hygiene; meat tenderizer; ss.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT misc_RNA 28

XX FT /*tag= a
 /*note= "ribonucleotide"

XX PN WO9849346-A1.

XX PD 05-NOV-1998.

XX PP 29-APR-1998; 98WO-US008677.

XX PR 29-APR-1997; 97US-0045228P.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Joyce GF, Breaker RR;

XX DR WPI; 1999-034670/03.

XX PT New catalytic DNA molecules - having site-specific endonuclease activity
 in a substrate nucleic acid, used for cleaving target nucleic acid
 sequences.

XX PS Disclosure; Page 59; 161pp; English.

XX CC This sequence is used in a method which involves the production of
 CC catalytic DNA molecules which can be used for cleaving target nucleic
 CC acid molecules. Such DNA molecules can be used in pharmaceutical and
 CC medical products (e.g. for wound debridement, clot dissolution), as well
 CC as in household items (e.g. detergents, dental hygiene products, meat
 CC tenderizers). Other suitable substrates include those comprising or
 CC produced by picornaviruses, hepadnaviridae, (e.g. HBV, HCV),
 CC papillomaviruses (e.g. HPV), gammaherpesvirinae (e.g. EBV),
 CC lymphocryptoviruses, leukemia viruses (e.g. HIV-1 and -11,
 CC flaviviruses, togaviruses, herpesviruses (including alphaherpesvirus and
 CC betaherpesviruses), cytomegaloviruses (CMV), influenza viruses, viruses
 CC and retroviruses contributing to immunodeficiency diseases and syndromes
 CC (e.g. HIV-1 and -2), simian and feline immunodeficiency viruses and
 CC bovine leukemia viruses. They can also be used as regulators of gene
 CC expression

XX Sequence 43 BP; 15 A; 8 C; 12 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 2; Length 43;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 DB 19 ACTCACTATAGGAGAGATG 38

RESULT 13
AAV82932
ID AAV82932 standard; DNA; 43 BP.
XX
AC AAV82932;
XX
DT 05-MAR-1999 (first entry)
XX
DE Enzymatic DNA primer 2.
XX
KW Enzyme; catalysis; cleavage; target; pharmaceutical; medical; substrate;
KW regulator; detergent; dental hygiene; meat tenderizer; primer; ss.
XX
OS Synthetic.
XX
PN WO9849346-A1.
XX
PD 05-NOV-1998.
XX
PF 29-APR-1998; 98WO-US008677.
XX
PR 29-APR-1997; 97US-0045228P.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Joyce GF, Breaker RR;
XX
PI WPI; 1999-034670/03.
XX
DR New catalytic DNA molecules - having site-specific endonuclease activity
PT in a substrate nucleic acid, used for cleaving target nucleic acid
PT sequences.
XX
PS Disclosure; Page 62; 16pp; English.
XX
CC This sequence is used in a method which involves the production of
CC catalytic DNA molecules which can be used for cleaving target nucleic
CC acid molecules. Such DNA molecules can be used in pharmaceutical and
CC medical products (e.g. for wound debridement, clot dissolution), as well
CC as in household items (e.g. detergents, dental hygiene products, meat
CC tenderisers). Other suitable substrates include those comprising or
CC produced by picornaviruses, hepadnaviridae (e.g. HBV, HCV),
CC papillomaviruses (e.g. HPV), gammaherpesvirinae (e.g. EBV, HHV-8),
CC lymphocryptoviruses, leukemia viruses (e.g. HTLV-1 and -II),
CC flaviviruses, togaviruses, herpesviruses (including alphaherpesviruses and
CC betaherpesviruses), cytomegaloviruses (CMV), influenza viruses, viruses
CC and retroviruses contributing to immunodeficiency diseases and syndromes
CC (e.g. HIV-1 and -2), simian and feline immunodeficiency viruses and
CC bovine leukemia viruses. They can also be used as regulators of gene
CC expression
XX
SQ Sequence 43 BP; 15 A; 9 C; 10 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 39
RESULT 14
AAV82932
ID AAV82932 standard; DNA; 43 BP.
XX
AC AAV82932;
XX
DT 09-JAN-2001 (first entry)
XX
DE Primer oligonucleotide sequence SEQ ID NO:11.
XX
KW Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;

KW substrate; template; primer; cleavage; hydrolytic cleavage;
KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.
XX
OS Synthetic.
XX
PN US6110462-A.
XX
PD 29-AUG-2000.
XX
PF 03-MAR-1999; 99US-00262142.
XX
PR 03-MAR-1999; 99US-00262142.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Barbas CF, Kandasamy S, Joyce G, Santoro SW;
XX
PI WPI; 2000-593449/56.
XX
DR Enzymatic DNA molecules containing modified nucleotides useful for
PT cleavage of RNA at a specified position.
XX
PS Example 2; Col 29; 52pp; English.
XX
CC The present invention describes a catalytic DNA molecule that
CC specifically cleaves a substrate nucleic acid at a defined cleavage site,
CC where the catalytic DNA molecule comprises at least one pyrimidine
CC nucleotide (1). (1) is a catalytic DNA molecule, capable of hydrolytic
CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule
CC is used to cleave RNA of almost any sequence. This is useful for cleavage
CC of single-stranded nucleic acid in the absence of a restriction
CC endonuclease site at a specific position. This is useful in methods of
CC e.g. cloning and genetic engineering. The catalytic core of the minimised
CC enzyme is composed of only 12 residues, making this one of the smallest
CC nucleic acid catalysts known. The catalytic core forms a compact hairpin
CC structure displaying the 3 imidazole-containing residues. The enzyme can
CC be made to cleave RNA of almost any sequence by simple alteration of the
CC two substrate-recognition domains that surround the catalytic core. The
CC enzyme operates with multiple turnover in the presence of micromolar
CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic
CC rate enhancement of approximately 1000000-fold compared to the
CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the
CC substrate-recognition properties of nucleic acid enzymes and the chemical
CC functionality of protein enzymes in a molecule that is small in size, yet
CC versatile and catalytically efficient. The present sequence represents an
CC oligonucleotide sequence which is used in the exemplification of the
CC present invention
XX
SQ Sequence 43 BP; 15 A; 9 C; 10 G; 9 T; 0 U; 0 Other;
Query Match 100.0%; Score 20; DB 3; Length 43;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAGAGATG 20
DB 19 ACTCACTATAGGAGAGATG 38
RESULT 15
AAV82934
ID AAV82934 standard; DNA; 43 BP.
XX
AC AAV82934;
XX
DT 09-JAN-2001 (first entry)
XX
DE Substrate oligonucleotide sequence SEQ ID NO:13.
XX
KW Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;
KW substrate; template; primer; cleavage; hydrolytic cleavage;
KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.
XX

OS Synthetic.
 PH Key Location/Qualifiers
 FT modified_base 28
 FT /tag= a
 FT /note= "adenosine ribonucleotide"

XX US6110462-A.
 XX 29-AUG-2000.
 XX 03-MAR-1999; 99US-00262142.
 XX 03-MAR-1999; 99US-00262142.
 XX (SCRI) SCRIPPS RES INST.
 XX Barbas CP, Kandaamy S, Joyce G, Santoro SW;
 XX WPI; 2000-593449/56.
 XX Enzymatic DNA molecules containing modified nucleotides useful for
 XX cleavage of RNA at a specified position.
 XX Example 1; Col 28; 52pp; English.

CC The present invention describes a catalytic DNA molecule that
 CC specifically cleaves a substrate nucleic acid at a defined cleavage site,
 CC where the catalytic DNA molecule comprises at least one pyrimidine
 CC nucleotide (1). (1) is a catalytic DNA molecule, capable of hydrolytic
 CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule
 CC is used to cleave RNA of almost any sequence. This is useful for cleavage
 CC of single-stranded nucleic acid in the absence of a restriction
 CC endonuclease site at a specific position. This is useful in methods of
 CC e.g. cloning and genetic engineering. The catalytic core of the minimised
 CC nucleic acid catalyst known. The catalytic core forms a compact hairpin
 CC structure displaying the 3 imidazole-containing residues. The enzyme can
 CC be made to cleave RNA of almost any sequence by simple alteration of the
 CC two substrate-recognition domains that surround the catalytic core. The
 CC enzyme operates with multiple turnover in the presence of micromolar
 CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic
 CC rate enhancement of approximately 1000000-fold compared to the
 CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the
 CC substrate-recognition properties of nucleic acid enzymes and the chemical
 CC functionality of protein enzymes in a molecule that is small in size, yet
 CC versatile and catalytically efficient. The present sequence represents an
 CC oligonucleotide sequence which is used in the exemplification of the
 CC present invention

XX Sequence 43 BP; 15 A; 8 C; 12 G; 8 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 3; Length 43;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 19 ACTCACTATAGGAGAGATG 38

RESULT 16
 AAL45393
 ID AAL45393 standard; DNA; 43 BP.
 AC AAL45393;
 XX 06-JUN-2002 (first entry)
 XX Zinc-dependent deoxyribozyme PCR primer p2.

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 19 ACTCACTATAGGAGAGATG 38

RESULT 16
 AAL45393
 ID AAL45393 standard; DNA; 43 BP.
 AC AAL45393;
 XX 06-JUN-2002 (first entry)
 XX Ion-dependent deoxyribozyme PCR primer p2.

XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
 KW Quencher; fluorophore; photodetector; ion concentration; ribozyme; PCR;

KN Primer; aa.
 XX Unidentified.
 XX WO200200006-A2.
 XX 03-JAN-2002.
 XX 27-JUN-2001; 2001WO-US020557.
 XX 27-JUN-2000; 2000US-00605558.
 XX (UNII) UNIV ILLINOIS FOUND.
 XX Lu Y, Li J;
 XX WPI; 2002-130823/17.
 XX New nucleic acid enzyme biosensors, useful for the sensitive and
 XX selective detection of ions, particularly metal ions e.g. lead ions, and
 XX for determining the concentration of a particular ion in a solution.
 XX Example 1; Page 23; 57pp; English.

CC The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependent on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion, particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a PCR primer
 CC used to isolate a ribozyme which may be used as a biosensor of the
 CC invention

XX Sequence 43 BP; 15 A; 9 C; 10 G; 9 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 6; Length 43;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 19 ACTCACTATAGGAGAGATG 38

RESULT 17
 AAL45391
 ID AAL45391 standard; DNA; 43 BP.
 XX AAL45391;
 XX 06-JUN-2002 (first entry)
 XX Zinc-dependent deoxyribozyme SEQ ID NO: 13.

XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
 KW Quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.
 XX Unidentified.
 XX WO200200006-A2.
 XX 03-JAN-2002.
 XX 27-JUN-2001; 2001WO-US020557.
 XX 27-JUN-2000; 2000US-00605558.
 XX (UNII) UNIV ILLINOIS FOUND.
 XX Lu Y, Li J;
 XX WPI; 2002-130823/17.

```

XX New nucleic acid enzyme biosensors, useful for the sensitive and
XX selective detection of ions, particularly metal ions e.g. lead ions, and
XX for determining the concentration of a particular ion in a solution.
XX
XX Example 1; Fig 2; 57pp; English.
XX
XX The present invention relates to biosensors, comprising a nucleic acid
XX enzyme dependent on an ion to produce a product, a quencher or/and a
XX fluorophore and a photodetector. The biosensors are useful for the
XX sensitive and selective detection of ions. The biosensors are useful in
XX methods of detecting the presence of an ion, particularly metal ions such
XX as lead. The biosensors may also be used to determine the concentration
XX of a particular ion in a solution. The present sequence is a ribozyme
XX which may be used as a biosensor of the invention
XX
XX Sequence 43 BP; 15 A; 9 C; 10 G; 9 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 6; Length 43;
XX Best Local Similarity 100.0%; Pred. No. 3.5;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ACTCACTATAGGAGAGATG 20
XX Db 19 ACTCACTATAGGAGAGATG 38
XX
XX RESULT 18
XX AAV36554
XX ID AAV36554 standard; DNA; 69 BP.
XX AC AAV36554;
XX
XX 09-OCT-1998 (first entry)
XX
XX Self-cleaving DNA of the invention.
XX
XX Self-cleaving DNA; allosteric DNA polynucleotide; modify; function;
XX configuration; detection; biosensor; 88.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX stem_loop 13..50
XX misc_structure 13..18
XX stem_loop 22..41
XX misc_structure 22..25
XX misc_structure 36..41
XX misc_structure 45..50
XX stem_loop 57..67
XX misc_structure 57..60
XX misc_structure 64..67
XX
XX W09827104-A1.
XX
XX 25-JUN-1998.
XX
XX 18-OEC-1997; 97WO-US024158.

```

```

XX 19-DEC-1996; 96US-0033684P.
XX 08-AUG-1997; 97US-005039P.
XX (UYIA ) UNIV YALB.
XX Breaker RR;
XX MPI; 1998-362715/31.
XX Allosteric polynucleotides - having modified function or configuration,
XX used for the production of biosensors.
XX Claim 6; Fig 14A; 106pp; English.
XX
XX The present sequence represents a self-cleaving DNA molecule of the
XX invention. The specification describes an allosteric DNA polynucleotide
XX which can modify a function or configuration of the polynucleotide with a
XX chemical effector and/or a physical signal. This polynucleotide can
XX comprise the present sequence. The allosteric polynucleotides can be used
XX for detecting the presence or concentrations of compounds such as amino
XX acids, peptides, nucleosides, nucleotides, steroids, microbial or
XX cellular metabolites, blood or urine components, pharmaceuticals,
XX pesticides, herbicides, or food toxins. The allosteric polynucleotides
XX can also be used for detecting physical signals such as radiation and
XX temperature changes. The polynucleotides can also be used in biosensors,
XX in which they are more stable than proteins
XX
XX Sequence 69 BP; 21 A; 13 C; 16 G; 19 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 20; DB 2; Length 69;
XX Best Local Similarity 100.0%; Pred. No. 3.6;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 ACTCACTATAGGAGAGATG 20
XX Db 14 ACTCACTATAGGAGAGATG 33
XX
XX RESULT 19
XX AAV36553
XX ID AAV36553 standard; DNA; 87 BP.
XX AC AAV36553;
XX
XX 09-OCT-1998 (first entry)
XX
XX Self-cleaving CAI variant DNA sequence.
XX
XX Self-cleaving DNA; allosteric DNA polynucleotide; modify; function;
XX configuration; detection; biosensor; 88.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX misc_structure 4..8
XX stem_loop 18..43
XX misc_structure 18..22
XX misc_structure 39..43
XX stem_loop 51..65
XX misc_structure 51..52
XX misc_structure 54..65
XX

```

```

FT  /note= "hybridises with nucleotides 51-52"
FT  66. .81
FT  /tag= b
FT  misc_structure
FT  66. .72
FT  /tag= i
FT  /note= "hybridises with nucleotides 75-81"
FT  75. .81
FT  /tag= j
FT  /note= "hybridises with nucleotides 65-72"
FT  83. .87
FT  /tag= k
FT  /note= "hybridises with nucleotides 4-8"
FT  WO9827104-A1.
FT  25-JUN-1998.
FT  18-DEC-1997; 97WO-US0241158.
FT  13-DEC-1996; 96US-0033684P.
FT  08-AUG-1997; 97US-0355039P.
FT  (UYIA ) UNIV YALE.
FT  Breaker RR;
FT  WPI; 1998-362715/31.
FT  Allosteric polynucleotides - having modified function or configuration,
FT  used for the production of biosensors.
FT  Claim 6; Fig 12B; 106pp; English.
FT  The present sequence represents a self-cleaving DNA molecule of the
FT  invention. The specification describes an allosteric DNA polynucleotide
FT  which can modify a function or configuration of the polynucleotide with a
FT  chemical effector and/or a physical signal. This polynucleotide can
FT  comprise the present sequence. The allosteric polynucleotides can be used
FT  for detecting the presence or concentrations of compounds such as amino
FT  acids, peptides, nucleosides, nucleotides, steroids, microbial or
FT  cellular metabolites, blood or urine components, pharmaceuticals,
FT  pesticides, herbicides, or food toxins. The allosteric polynucleotides
FT  can also be used for detecting physical signals such as radiation and
FT  temperature changes. The polynucleotides can also be used in biosensors,
FT  in which they are more stable than proteins
FT  Sequence 87 BP; 28 A; 14 C; 26 G; 19 T; 0 U; 0 Other;
FT  Query Match 100.0%; Score 20; DB 2; Length 87;
FT  Best Local Similarity 100.0%; Pred. No. 3.6;
FT  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT  QY 1 ACTCACTATAGGAGAGATG 20
FT  14 ACTCACTATAGGAGAGATG 33
FT  Db
FT  RESULT 20
FT  AAL45390
FT  ID AAL45390 standard; DNA; 97 BP.
FT  AC AAL45390;
FT  DT 06-JUN-2002 (first entry)
FT  DE RNA-cleaving deoxyribozyme #1.
FT  KW Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
FT  quenchers; fluorophore; photodetector; ion concentration; ribozyme; ss.
FT  OS Unidentified.
FT  Key Location/Qualifiers

```

```

FT  misc_binding 11. .14
FT  /tag= a
FT  /bound_moiety= "nucleotides 82-79"
FT  /note= "binds nucleotides 82-79 of itself"
FT  23. .38
FT  stem_loop
FT  /tag= b
FT  misc_binding 79. .82
FT  /tag= c
FT  /bound_moiety= "nucleotides 14-11"
FT  /note= "binds nucleotides 14-11 of itself"
FT  WO200200006-A2.
FT  03-JAN-2002.
FT  27-JUN-2001; 2001WO-US020557.
FT  27-JUN-2000; 2000US-00605558.
FT  (UNII ) UNIV ILLINOIS FOUND.
FT  Lu Y, Li J;
FT  WPI; 2002-130823/17.
FT  New nucleic acid enzyme biosensors, useful for the sensitive and
FT  selective detection of ions, particularly metal ions e.g. lead ions, and
FT  for determining the concentration of a particular ion in a solution.
FT  Example 1; Fig 1; 57pp; English.
FT  The present invention relates to biosensors, comprising a nucleic acid
FT  enzyme dependent on an ion to produce a product, a quencher or/and a
FT  fluorophore and a photodetector. The biosensors are useful for the
FT  sensitive and selective detection of ions. The biosensors are useful in
FT  methods of detecting the presence of an ion, particularly metal ions such
FT  as lead. The biosensors may also be used to determine the concentration
FT  of a particular ion in a solution. The present sequence is a ribozyme
FT  which may be used as a biosensor of the invention
FT  Sequence 97 BP; 17 A; 13 C; 15 G; 12 T; 0 U; 40 Other;
FT  Query Match 100.0%; Score 20; DB 6; Length 97;
FT  Best Local Similarity 100.0%; Pred. No. 3.6;
FT  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT  QY 1 ACTCACTATAGGAGAGATG 20
FT  9 ACTCACTATAGGAGAGATG 28
FT  Db
FT  RESULT 21
FT  AAL45385/c
FT  ID AAL45385 standard; DNA; 97 BP.
FT  AC AAL45385;
FT  DT 06-JUN-2002 (first entry)
FT  DE Ion-dependent deoxyribozyme random DNA pool template.
FT  KW Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
FT  quenchers; fluorophore; photodetector; ion concentration; ribozyme; ss.
FT  OS Unidentified.
FT  WO200200006-A2.
FT  03-JAN-2002.
FT  27-JUN-2001; 2001WO-US020557.
FT  27-JUN-2000; 2000US-00605558.

```

XX (UNIT) UNIV ILLINOIS FOUND.
 PA Lu Y, Li J;
 PI WPI; 2002-130823/17.
 XX
 DR New nucleic acid enzyme biosensors, useful for the sensitive and
 XX selective detection of ions, particularly metal ions e.g. lead ions, and
 PT for determining the concentration of a particular ion in a solution.
 PT
 XX
 PS Example 1; Page 23; 57pp; English.
 XX
 CC The present invention relates to biosensors, comprising a nucleic acid
 CC enzyme dependent on an ion to produce a product, a quencher or/and a
 CC fluorophore and a photodetector. The biosensors are useful for the
 CC sensitive and selective detection of ions. The biosensors are useful in
 CC methods of detecting the presence of an ion. Particularly metal ions such
 CC as lead. The biosensors may also be used to determine the concentration
 CC of a particular ion in a solution. The present sequence is a sequence
 CC described in the exemplification of the invention
 XX
 SQ Sequence 97 BP; 12 A; 16 C; 12 G; 17 T; 0 U; 40 Other;
 Query Match 100.0%; Score 20; DB 6; Length 97;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTCCTATAGGAGAGATG 20
 DB 89 ACTCCTATAGGAGAGATG 70

RESULT 22
 AAT34901
 ID AAT34901 standard; DNA; 107 BP.
 AC AAT34901;
 XX
 DT 11-FEB-1997 (first entry)
 DE Single stranded DNA with site specific endonuclease activity.
 KW Endonuclease; cleavage; RNA; catalyst; wound debridement;
 KW clot dissolution; meat tenderiser; viral infection; dental hygiene;
 KW detergent; enzyme isolation; ss.
 OS Synthetic.
 PH Key Location/Qualifiers
 FT misc_feature 28
 FT /*tag= a
 FT /note= "adenosine ribonucleotide"
 PN WO9617086-A1.
 XX
 PD 06-JUN-1996.
 PP 01-DEC-1995; 95WO-US015580.
 PR 02-DEC-1994; 94US-00349023.
 PR 07-JUN-1995; 95US-00472194.
 XX
 PA (SCRI) SCRIPPS RBS INST.
 PI Joyce GP, Breaker RR;
 XX WPI; 1996-286834/29.
 XX
 PT New DNA with site specific endonuclease activity - partic. for cleavage
 PT of RNA, e.g. for medical use.
 PS Claim 11; Page 66; 114pp; English.

XX AAT34891-T34967 are single stranded (ss), non-naturally occurring DNA
 CC molecules that have site-specific endonuclease activity. The catalytic
 CC DNA molecules specifically cleave ss nucleic acids, esp. RNA sequences.
 CC The catalytic DNAs may be used in medicine e.g. for wound debridement,
 CC clot dissolution, etc., or in detergents, dental hygiene products and
 CC meat tenderisers. The DNAs may be useful to treat viral infections such
 CC as HIV when targetted to viral nucleic acid and they may be expressed in
 CC a target host cell. In vitro selection of the DNAs allows the isolation
 CC of catalysts/enzymes without the need for prior knowledge of their
 CC compsn. or structure
 XX
 SQ Sequence 107 BP; 20 A; 14 C; 19 G; 14 T; 0 U; 40 Other;
 Query Match 100.0%; Score 20; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTCCTATAGGAGAGATG 20
 DB 19 ACTCCTATAGGAGAGATG 38

RESULT 23
 AAV82944
 ID AAV82944 standard; DNA; 107 BP.
 AC AAV82944;
 XX
 DT 05-MAR-1999 (first entry)
 DE Enzymatic DNA 107mer oligomer.
 XX
 KW Enzyme; catalysis; cleavage; target; pharmaceutical; medical; substrate;
 KW regulator; detergent; dental hygiene; meat tenderiser; ss.
 OS Synthetic.
 PH Key Location/Qualifiers
 FT misc_RNA 28
 FT /*tag= a
 FT /note= "ribonucleotide"
 PN WO9849346-A1.
 XX
 PD 05-NOV-1998.
 XX
 PF 29-APR-1998; 98WO-US008677.
 XX
 PR 29-APR-1997; 97US-0045228P.
 XX
 PA (SCRI) SCRIPPS RBS INST.
 XX
 PI Joyce GP, Breaker RR;
 XX
 DR WPI; 1999-034670/03.
 XX
 PT New catalytic DNA molecules - having site-specific endonuclease activity
 PT in a substrate nucleic acid, used for cleaving target nucleic acid
 PT sequences.
 XX
 PS Disclosure; Page 70; 161pp; English.
 XX
 CC This sequence is used in a method which involves the production of
 CC catalytic DNA molecules which can be used for cleaving target nucleic
 CC acid molecules. Such DNA molecules can be used in pharmaceutical and
 CC medical products (e.g. for wound debridement, clot dissolution), as well
 CC as in household items (e.g. detergents, dental hygiene products, meat
 CC tenderisers). Other suitable substrates include those comprising or
 CC produced by picornaviruses, hepadnaviridae, (e.g. HBV, HCV),
 CC papillomaviruses (e.g. HPV), gammaherpesvirinae (e.g. EBV),
 CC lymphocryptoviruses, leukemia viruses (e.g. HTLV-1 and -11),
 CC flaviviruses, togaviruses, herpesviruses (including alphaherpesviruses and

CC betaherpesviruses), cytomegaloviruses (CMV), influenza viruses, viruses
 CC and retroviruses contributing to immunodeficiency diseases and syndromes
 CC (e.g. HIV-1 and -2), simian and feline immunodeficiency viruses and
 CC bovine leukemia viruses. They can also be used as regulators of gene
 CC expression

XX Sequence 107 BP; 20 A; 14 C; 19 G; 14 T; 0 U; 40 Other;
 Query Match 100.0%; Score 20; DB 2; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ACTCACTATAGGAGAGATG 20
 |||||
 Db 19 ACTCACTATAGGAGAGATG 38

RESULT 24
 AAA92244
 ID AAA92244 standard; DNA; 107 BP.

XX AC AAA92244;

XX BT 09-JAN-2001 (first entry)

XX DE Synthetic oligomer sequence SEQ ID NO:23.

XX Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;
 KW substrate; primer; cleavage; hydrolytic cleavage;
 KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.

XX OS Synthetic.

XX FN US6110462-A.

XX FD 29-AUG-2000.

XX PF 03-MAR-1999; 99US-00262142.

XX PR 03-MAR-1999; 99US-00262142.

XX PA (SCRI) SCRIPPS RBS INST.

XX PI Barbas CP, Kandasamy S, Joyce G, Santoro SW;

XX WPI; 2000-593449/56.

XX Enzymatic DNA molecules containing modified nucleotides useful for
 cleavage of RNA at a specified position.

XX Example 4; Col 33; 52pp; English.

CC The present invention describes a catalytic DNA molecule that
 CC specifically cleaves a substrate nucleic acid at a defined cleavage site,
 CC where the catalytic DNA molecule comprises at least one pyrimidine
 CC nucleotide (I). (i) is a catalytic DNA molecule, capable of hydrolytic
 CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule
 CC is used to cleave RNA of almost any sequence. This is useful for cleavage
 CC of single-stranded nucleic acid in the absence of a restriction
 CC endonuclease site at a specific position. This is useful in methods of
 CC e.g. cloning and genetic engineering. The catalytic core of the minimised
 CC enzyme is composed of only 12 residues, making this one of the smallest
 CC nucleic acid catalysts known. The catalytic core forms a compact hairpin
 CC structure displaying the 3 imidazole-containing residues. The enzyme can
 CC be made to cleave RNA of almost any sequence by simple alteration of the
 CC two substrate-recognition domains that surround the catalytic core. The
 CC enzyme operates with multiple turnover in the presence of micromolar
 CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic
 CC rate enhancement of approximately 1000000-fold compared to the
 CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the
 CC substrate-recognition properties of nucleic acid enzymes and the chemical
 CC functionality of protein enzymes in a molecule that is small in size, yet
 CC versatile and catalytically efficient. The present sequence represents an

CC oligonucleotide sequence which is used in the exemplification of the
 CC present invention

XX Sequence 107 BP; 20 A; 14 C; 19 G; 14 T; 0 U; 40 Other;

Query Match 100.0%; Score 20; DB 3; Length 107;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ACTCACTATAGGAGAGATG 20
 |||||
 Db 19 ACTCACTATAGGAGAGATG 38

RESULT 25

AA45461
 ID AAL45461 standard; DNA; 107 BP.

XX AC AAL45461;

XX DT 06-JUN-2002 (first entry)

XX DB G3 deoxyribose.

XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
 KW quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.

XX OS Unidentified.

XX Key Location/Qualifiers

FT misc_binding 1..2

FT /tag= a
 FT /bound_moiety= "nucleotides 108-107"
 FT /note= "binds nucleotides 108-107 of itself"

FT 3

FT misc_binding /tag= b
 FT /bound_moiety= "nucleotide 105"
 FT /note= "binds nucleotide 105 of itself"

FT 5

FT misc_binding /tag= c
 FT /bound_moiety= "nucleotide 104"
 FT /note= "binds nucleotide 104 of itself"

FT 7..8

FT misc_binding /tag= d
 FT /bound_moiety= "nucleotides 102-101"
 FT /note= "binds nucleotides 102-101 of itself"

FT 16..18

FT misc_binding /tag= e
 FT /bound_moiety= "nucleotides 69-67"
 FT /note= "binds nucleotides 69-67 of itself"

FT 22

FT misc_binding /tag= f
 FT /bound_moiety= "nucleotide 66"
 FT /note= "binds nucleotide 66 of itself"

FT 24

FT misc_binding /tag= g
 FT /bound_moiety= "nucleotide 64"
 FT /note= "binds nucleotide 64 of itself"

FT 28

FT misc_binding /tag= h
 FT /bound_moiety= "nucleotide 24"
 FT /note= "binds nucleotide 24 of itself"

FT 31..50

FT stem_loop /tag= i
 FT /bound_moiety= "nucleotide 22"
 FT /note= "binds nucleotide 22 of itself"

FT 53..63

FT stem_loop /tag= j
 FT /bound_moiety= "nucleotide 24"
 FT /note= "binds nucleotide 24 of itself"

FT 64

FT misc_binding /tag= k
 FT /bound_moiety= "nucleotide 24"
 FT /note= "binds nucleotide 24 of itself"

FT 66

FT misc_binding /tag= l
 FT /bound_moiety= "nucleotide 22"
 FT /note= "binds nucleotide 22 of itself"

FT 67..69

```

PT /tag= m
PT /bound_moiety= "nucleotides 18-16"
PT /note= "binds nucleotides 18-16 of itself"
PT 76..100
PT /tag= n
PT /tag= o
PT /bound_moiety= "nucleotides 8-7"
PT /note= "binds nucleotides 8-7 of itself"
PT 104
PT /tag= p
PT /bound_moiety= "nucleotide 5"
PT /note= "binds nucleotide 5 of itself"
PT 105
PT /tag= q
PT /bound_moiety= "nucleotide 3"
PT /note= "binds nucleotide 3 of itself"
PT 107..108
PT /tag= r
PT /bound_moiety= "nucleotides 2-1"
PT /note= "binds nucleotides 2-1 of itself"

```

MO200200006-A2.

03-JAN-2002.

27-JUN-2001; 2001MO-US020557.

27-JUN-2000; 2000US-00605558.

(UNII) UNIV ILLINOIS FOUNO.

Lu Y, Li J;

WPI; 2002-130823/17.

New nucleic acid enzyme biosensors, useful for the sensitive and selective detection of ions, particularly metal ions e.g. lead ions, and for determining the concentration of a particular ion in a solution.

Example 1; Fig 7; 57pp; English.

The present invention relates to biosensors, comprising a nucleic acid enzyme dependent on an ion to produce a product, a quencher or/and a fluorophore and a photodetector. The biosensors are useful for the sensitive and selective detection of ions. The biosensors are useful in methods of detecting the presence of an ion, particularly metal ions such as lead. The biosensors may also be used to determine the concentration of a particular ion in a solution. The present sequence is a ribozyme which may be used as a biosensor of the invention

Sequence 107 BP; 27 A; 20 C; 26 G; 19 T; 0 U; 15 Other;

Query Match 100.0%; Score 20; DB 6; Length 107;
Best Local Similarity 100.0%; Pred. NO. 3.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
19 ACTCACTATAGGAGAGATG 38

RESULT 26

ABK10821

ID ABK10821 standard; DNA; 123 BP.

AC ABK10821;

DT 21-MAY-2002 (first entry)

DE Self-capping deoxyribozyme selection, acceptor construct.

KW Self-phosphorylating deoxyribozyme; conjoined polynucleotide; catalyst;

ribozyme; self-capping deoxyribozyme; ss.

Synthetic.

Key Location/Qualifiers

misc_binding 6..38

misc_binding /tag= a

misc_binding /bound_moiety= "ABK10823"

misc_binding /note= "Forms a double stranded region with the template strand shown in ABK10823"

misc_feature 23..24

misc_feature /tag= b

misc_feature /note= "Capping and ligation site of deoxyribozyme"

misc_feature 39..108

misc_feature /tag= c

misc_feature /note= "Random sequence of 70 nucleotides"

primer_bind 109..123

primer_bind /tag= d

primer_bind /bound_moiety= "Primer 1"

primer_bind /note= "The sequence for primer 1 is not given in the specification"

MO200129249-A2.

26-APR-2001.

13-OCT-2000; 2000MO-US028508.

15-OCT-1999; 99US-0159808P.

(UYIA) UNIV YALE.

Breaker RR;

WPI; 2002-226120/28.

New conjoined polynucleotides comprising a kinase domain and an adenylyase domain, useful in DNA cloning, for carrying out sequential polynucleotide manipulations in a serial fashion, or for polynucleotide manipulations.

Example 1; Fig 5B; 55pp; English.

The invention describes a conjoined polynucleotide comprising at least two catalytic domains which function in concert to provide a chemical transformation involving multiple sequential reactions. The use of the conjoined polynucleotides as catalysts offer advantages over protein-based enzymes in cloning and in a number of commercial and industrial processes. Conjoined polynucleotides made from DNA are expected to be much more stable, and can be easily made by automated oligonucleotide synthesis, and DNA is significantly more resistant to hydrolytic degradation compared to RNA. Conjoined DNA and RNA may be selected for their ability to function on solid support and are expected to retain their activity when immobilised. This sequence represents an acceptor construct used in the selection of self-capping deoxyribozymes, described in the invention

Sequence 123 BP; 17 A; 10 C; 14 G; 12 T; 0 U; 70 Other;

Query Match 100.0%; Score 20; DB 6; Length 123;

Best Local Similarity 100.0%; Pred. NO. 3.7;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20

14 ACTCACTATAGGAGAGATG 33

RESULT 27

ABK10822

ID ABK10822 standard; DNA; 123 BP.

AC ABK10822;

XX

```

DT 21-MAY-2002 (first entry)
XX Self-phosphorylating deoxyribozyme selection, acceptor construct.
XX Self-phosphorylating deoxyribozyme, conjoined polynucleotide, catalyst;
KW ribozyme; ss.
XX Synthetic.
XX Key Location/Qualifiers
FH misc_binding 6..38
FT /tag= a
FT /bound moiety= "ABK10807"
FT /note= "Forms a double stranded region with the template
FT strand shown in ABK10807."
FT misc_feature 23..24
FT /tag= b
FT /note= "Phosphorylation and ligation site"
FT misc_feature 39..108
FT /tag= c
FT /note= "Random sequence of 70 nucleotides"
FT primer_bind 139..123
FT /tag= d
FT /bound moiety= "Primer 1"
FT /note= "The sequence for primer 1 is not given in the
FT specification"
XX WO200129249-A2.
XX 26-APR-2001.
XX 13-OCT-2000; 2000WO-US028508.
XX 15-OCT-1999; 99US-0159808P.
XX (UYVA ) UNIV YALE.
XX Breaker RR;
XX MPI; 2002-226120/28.
XX New conjoined polynucleotides comprising a kinase domain and an adenylase
XX domain, useful in DNA cloning, for carrying out sequential polynucleotide
XX manipulations in a serial fashion, or for polynucleotide manipulations.
XX Example 1; Fig 4B; 55pp; English.
XX The invention describes a conjoined polynucleotide comprising at least
XX two catalytic domains which function in concert to provide a chemical
XX transformation involving multiple sequential reactions. The use of the
XX conjoined polynucleotides as catalysts offer advantages over protein-
XX based enzymes in cloning and in a number of commercial and industrial
XX processes. Conjoined polynucleotides made from DNA are expected to be
XX much more stable, and can be easily made by automated oligonucleotide
XX synthesis, and DNA is significantly more resistant to hydrolytic
XX degradation compared to RNA. Conjoined DNA and RNA may be selected for
XX their ability to function on solid support and are expected to retain
XX their activity when immobilised. This sequence represents an acceptor
XX construct used in the selection of self-phosphorylating deoxyribozymes,
XX described in the invention
XX Sequence 123 BP; 17 A; 9 C; 15 G; 12 T; 0 U; 70 Other;
XX Query Match 100.0%; Score 20; DB 6; Length 123;
XX Best Local Similarity, 100.0%; Pred. No. 3.7;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 ACTCACTATAGGAAGAGATG 20
XX Db 14 ACTCACTATAGGAAGAGATG 33
XX RESULT 28

```

```

AAL45458
ID AAL45458 standard; DNA; 19 BP.
XX AAL45458;
XX 06-JUN-2002 (first entry)
XX RNA-cleaving deoxyribozyme substrate #5.
XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
XX quencher; fluorophore; photodetector; ion concentration; ribozyme;
XX substrate; ss.
XX Unidentified.
XX Key Location/Qualifiers
FH misc_binding 1..8
FT /tag= a
FT /bound moiety= "RNA-cleaving deoxyribozyme"
FT /note= "Binds nucleotides 41-34 of the deoxyribozyme"
FT shown in AAL45457"
FT misc_RNA 9
FT /tag= b
FT misc_binding 12..19
FT /tag= c
FT /bound moiety= "RNA-cleaving deoxyribozyme"
FT /note= "Binds nucleotides 8-1 of the deoxyribozyme shown
FT in AAL45457"
XX WO200200006-A2.
XX 03-JAN-2002.
XX 27-JUN-2001; 2001WO-US020557.
XX 27-JUN-2000; 2000US-00605558.
XX (UNII ) UNIV ILLINOIS FOUND.
XX Lu Y, Li J;
XX MPI; 2002-130823/17.
XX New nucleic acid enzyme biosensors, useful for the sensitive and
XX selective detection of ions, particularly metal ions e.g. lead ions, and
XX for determining the concentration of a particular ion in a solution.
XX Example 1; Fig 6; 57pp; English.
XX The present invention relates to biosensors, comprising a nucleic acid
XX enzyme dependent on an ion to produce a product, a quencher or/and a
XX fluorophore and a photodetector. The biosensors are useful for the
XX sensitive and selective detection of ions. The biosensors are useful in
XX methods of detecting the presence of an ion, particularly metal ions such
XX as lead. The biosensors may also be used to determine the concentration
XX of a particular ion in a solution. The present sequence is a substrate of
XX a ribozyme which may be used as a biosensor of the invention
XX Sequence 19 BP; 7 A; 3 C; 5 G; 4 T; 0 U; 0 Other;
XX Query Match 95.0%; Score 19; DB 6; Length 19;
XX Best Local Similarity 100.0%; Pred. No. 10;
XX Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 2 CTCACCTATAGGAAGAGATG 20
XX Db 1 CTCACCTATAGGAAGAGATG 19
XX RESULT 29
XX AAV36569
XX ID AAV36569 standard; DNA; 65 BP.
XX

```



```

AC AAV36569;
DT 09-OCT-1998 (first entry)
DE Self-cleaving DNA (deoxyribozyme) sequence HD2.
KW Self-cleaving DNA; allosteric DNA polynucleotide; modify; function;
KW configuration; detection; biosensor; ss.
OS Synthetic.
XX
PH Key Location/Qualifiers
FT misc_feature 12..12
FT /tag= a
FT /note= "RNA A linkage"
XX
XX MO9827104-A1.
XX 25-JUN-1998.
XX 18-DEC-1997; 97WO-US024158.
XX 19-DEC-1996; 96US-0033684P.
XX 08-AUG-1997; 97US-0055039P.
XX (UYVA ) UNIV VALE.
XX Breaker RR;
XX WPI, 1998-362715/31.
XX
XX Allosteric polynucleotides - having modified function or configuration,
XX used for the production of biosensors.
XX Claim 6; Fig 19A; 106pp; English.
XX
XX The present sequence represents a self-cleaving DNA molecule HD2
XX (ribozyme) of the invention. The specification describes an allosteric
XX DNA polynucleotide which can modify a function or configuration of the
XX polynucleotide with a chemical effector and/or a physical signal. This
XX polynucleotide can comprise the present sequence. The allosteric
XX polynucleotides can be used for detecting the presence or concentrations
XX of compounds such as amino acids, peptides, nucleosides, nucleotides,
XX steroids, microbial or cellular metabolites, blood or urine components,
XX pharmaceuticals, pesticides, herbicides, or food toxins. The allosteric
XX polynucleotides can also be used for detecting physical signals such as
XX radiation and temperature changes. The polynucleotides can also be used
XX in biosensors, in which they are more stable than proteins
XX
XX Sequence 65 BP; 16 A; 10 C; 23 G; 15 T; 0 U; 1 Other;
XX
Query Match 95.0%; Score 19; DB 2; Length 65;
Best Local Similarity 95.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 1; Indels 3; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 3 ACTCACTATAGGAGAGATG 22

RESULT 30
AAL45402
IC AAL45402 standard; DNA; 43 BP.
XX
XX AAL45402;
AC
XX 06-JUN-2002 (first entry)
DT
DE Cobalt-dependent deoxyribozyme SEQ ID NO: 24.
XX
XX Biosensor; ion detection; lead; zinc; cobalt; nucleic acid enzyme;
XX quencher; fluorophore; photodetector; ion concentration; ribozyme; ss.
XX

```

```

OS Unidentified.
XX
XX WO200200006-A2.
XX
XX 03-JAN-2002.
XX
XX 27-JUN-2001; 2001WO-US020557.
XX
XX 27-JUN-2000; 2000US-00605558.
XX
XX (UNII ) UNIV ILLINOIS FOUND.
XX
XX Lu Y, Li J;
XX
XX WPI, 2002-130823/17.
XX
XX New nucleic acid enzyme biosensors, useful for the sensitive and
XX selective detection of ions, particularly metal ions e.g. lead ions, and
XX for determining the concentration of a particular ion in a solution.
XX
XX Example 1; Fig 3; 57pp; English.
XX
XX The present invention relates to biosensors, comprising a nucleic acid
XX enzyme dependent on an ion to produce a product, a quencher or/and a
XX fluorophore and a photodetector. The biosensors are useful for the
XX sensitive and selective detection of ions. The biosensors are useful in
XX methods of detecting the presence of an ion, particularly metal ions such
XX as lead. The biosensors may also be used to determine the concentration
XX of a particular ion in a solution. The present sequence is a ribozyme
XX which may be used as a biosensor of the invention
XX
XX Sequence 43 BP; 15 A; 9 C; 11 G; 8 T; 0 U; 0 Other;
XX
Query Match 92.0%; Score 18.4; DB 6; Length 43;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACTCACTATAGGAGAGATG 20
Db 19 ACCCACTATAGGAGAGATG 38

RESULT 31
AAV82933
ID AAV82933 standard; DNA, 19 BP.
XX
XX AAV82933;
XX
XX 05-MAR-1999 (first entry)
XX
XX Enzymatic DNA 19mer substrate.
XX
XX Enzyme; catalysis; cleavage; target; pharmaceutical; medical; substrate;
XX regulator; detergent; dental hygiene; meat tenderizer; ss.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX misc_RNA 8
XX /tag= a
XX /note= "ribonucleotide"
XX
XX WO9849346-A1.
XX
XX 05-NOV-1998.
XX
XX 29-APR-1998; 98WO-US008677.
XX
XX 29-APR-1997; 97US-0045228P.
XX
XX (SCHI ) SCRIpps RES INST.
XX
XX Joyce GP, Breaker RR;
XX

```

XX DR WPI; 1999-034670/03.

XX PT New catalytic DNA molecules - having site-specific endonuclease activity

XX PT in a substrate nucleic acid, used for cleaving target nucleic acid

XX PT sequences.

XX PS Disclosure; Page 66; 161pp; English.

XX CC This sequence is used in a method which involves the production of

XX CC catalytic DNA molecules which can be used for cleaving target nucleic

XX CC acid molecules. Such DNA molecules can be used in pharmaceutical and

XX CC medical products (e.g. for wound debridement, clot dissolution), as well

XX CC as in household items (e.g. detergents, dental hygiene products, meat

XX CC tenderizers). Other suitable substrates include those comprising or

XX CC produced by picornaviruses, hepadnaviridae, (e.g. HBV, HCV),

XX CC papillomaviruses (e.g. HPV), gammaherpesvirinae (e.g. EBV),

XX CC lymphocryptoviruses, leukemia viruses (e.g. HTLV-I and -II),

XX CC flaviviruses, togaviruses, herpesviruses (including alphaherpesviruses and

XX CC betaherpesviruses), cytomagaloviruses (CMV), influenza viruses, viruses

XX CC and retroviruses contributing to immunodeficiency diseases and syndromes

XX CC (e.g. HIV-1 and -2), simian and feline immunodeficiency viruses and

XX CC bovine leukemia viruses. They can also be used as regulators of gene

XX CC expression

XX SQ Sequence 19 BP; 7 A; 2 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 90.0%; Score 18; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCACATAGGAGAGATG 20

Db 1 TCACATAGGAGAGATG 18

RESULT 32

AAA92228

ID AAA92228 standard; DNA; 19 BP.

XX AC AAA92228;

XX DT 09-JAN-2001 (first entry)

XX DE Primer oligonucleotide sequence SEQ ID NO:7.

XX KW Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;

XX KW substrate; template; primer; cleavage; hydrolytic cleavage;

XX KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.

XX OS Synthetic.

XX PH Key Location/Qualifiers

XX FT modified_base /tag= a

XX FT /note= "adenosine ribonucleotide"

XX PN US6110462-A.

XX PD 29-AUG-2000.

XX PP 03-MAR-1999; 99US-00262142.

XX PR 03-MAR-1999; 99US-00262142.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Barbas CF, Kandasamy S, Joyce G, Santoro SM;

XX DR WPI; 2000-593449/56.

XX PT Enzymatic DNA molecules containing modified nucleotides useful for

XX PT cleavage of RNA at a specified position.

XX PS Example 2; Col 29; 52pp; English.

XX CC The present invention describes a catalytic DNA molecule that

XX CC specifically cleaves a substrate nucleic acid at a defined cleavage site,

XX CC where the catalytic DNA molecule comprises at least one pyrimidine

XX CC nucleotide (I). (II) is a catalytic DNA molecule, capable of hydrolytic

XX CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule

XX CC is used to cleave RNA of almost any sequence. This is useful for cleavage

XX CC of single-stranded nucleic acid in the absence of a restriction

XX CC endonuclease site at a specific position. This is useful in methods of

XX CC e.g. cloning and genetic engineering. The catalytic core of the minimised

XX CC enzyme is composed of only 12 residues, making this one of the smallest

XX CC nucleic acid catalysts known. The catalytic core forms a compact hairpin

XX CC structure displaying the 3 imidazole-containing residues. The enzyme can

XX CC be made to cleave RNA of almost any sequence by simple alteration of the

XX CC two substrate-recognition domains that surround the catalytic core. The

XX CC enzyme operates with multiple turnover in the presence of micromolar

XX CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic

XX CC rate enhancement of approximately 1000000-fold compared to the

XX CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the

XX CC substrate-recognition properties of nucleic acid enzymes and the chemical

XX CC functionality of protein enzymes in a molecule that is small in size, yet

XX CC versatile and catalytically efficient. The present sequence represents an

XX CC oligonucleotide sequence which is used in the exemplification of the

XX CC present invention

XX SQ Sequence 19 BP; 7 A; 2 C; 6 G; 4 T; 0 U; 0 Other;

Query Match 90.0%; Score 18; DB 3; Length 19;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TCACATAGGAGAGATG 20

Db 1 TCACATAGGAGAGATG 18

RESULT 33

AAA92223

ID AAA92223 standard; DNA; 19 BP.

XX AC AAA92223;

XX DT 09-JAN-2001 (first entry)

XX DE Termination oligonucleotide sequence SEQ ID NO:2.

XX KW Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;

XX KW substrate; template; primer; cleavage; hydrolytic cleavage;

XX KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.

XX OS Synthetic.

XX PH Key Location/Qualifiers

XX FT modified_base /tag= a

XX FT /note= "adenosine ribonucleotide"

XX PN US6110462-A.

XX PD 29-AUG-2000.

XX PP 03-MAR-1999; 99US-00262142.

XX PR 03-MAR-1999; 99US-00262142.

XX PA (SCRI) SCRIPPS RES INST.

XX PI Barbas CF, Kandasamy S, Joyce G, Santoro SM;

XX DR WPI; 2000-593449/56.

XX PT Enzymatic DNA molecules containing modified nucleotides useful for

XX PT cleavage of RNA at a specified position.

PT Enzymatic DNA molecules containing modified nucleotides useful for
 PT cleavage of RNA at a specified position.
 XX
 XX Example 1; Col 7; 52pp; English.
 XX
 CC The present invention describes a catalytic DNA molecule that
 CC specifically cleaves a substrate nucleic acid at a defined cleavage site,
 CC where the catalytic DNA molecule comprises at least one pyrimidine
 CC nucleotide (1). (1) is a catalytic DNA molecule, capable of hydrolytic
 CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule
 CC is used to cleave RNA of almost any sequence. This is useful for cleavage
 CC of single-stranded nucleic acid in the absence of a restriction
 CC endonuclease site at a specific position. This is useful in methods of
 CC e.g. cloning and genetic engineering. The catalytic core of the smallest
 CC nucleic acid catalysts known. The catalytic core forms a compact hairpin
 CC structure displaying the 3 imidazole-containing residues. The enzyme can
 CC be made to cleave RNA of almost any sequence by simple alteration of the
 CC two substrate-recognition domains that surround the catalytic core. The
 CC enzyme operates with multiple turnover in the presence of micromolar
 CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic
 CC rate enhancement of approximately 1000000-fold compared to the
 CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the
 CC substrate-recognition properties of nucleic acid enzymes and the chemical
 CC functionality of protein enzymes in a molecule that is small in size, yet
 CC versatile and catalytically efficient. The present sequence represents an
 CC oligonucleotide sequence which is used in the exemplification of the
 CC present invention
 XX
 XX Sequence 19 BP; 7 A; 2 C; 6 G; 4 T; 0 U; 0 Other;
 SQ
 Query Match 90.0%; Score 18; DB 3; Length 19;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 TCACCTATAGGAGAGATG 20
 DB 1 TCACCTATAGGAGAGATG 18
 RESULT 34
 AAA92233
 ID AAA92233 standard; DNA; 19 BP.
 XX
 AC AAA92233;
 XX
 DT 09-JAN-2001 (first entry)
 XX
 DE Substrate oligonucleotide sequence SEQ ID NO:12.
 XX
 KW Catalytic; enzymatic; modified nucleotide; enzyme; site-specific; target;
 KW substrate; template; primer; cleavage; hydrolytic cleavage;
 KW nucleic acid phosphoester bond; cloning; genetic engineering; ss.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH modified_base 8
 FT /*tag= a
 FT /note= "adenosine ribonucleotide"
 FT
 PT
 PT US6110462-A.
 PN
 XX 29-AUG-2000.
 PD
 XX
 XX 03-MAR-1999; 99US-00262142.
 PF
 XX 03-MAR-1999; 99US-00262142.
 PR
 XX (SRI) SCRIPPS RES INST.
 XX
 PA Barbas CP, Kandasamy S, Joyce G, Santoro SW;
 PI
 XX

DR WPI, 2000-593449/56.
 XX Enzymatic DNA molecules containing modified nucleotides useful for
 PT cleavage of RNA at a specified position.
 XX
 XX Example 3; Col 31; 52pp; English.
 XX
 CC The present invention describes a catalytic DNA molecule that
 CC specifically cleaves a substrate nucleic acid at a defined cleavage site,
 CC where the catalytic DNA molecule comprises at least one pyrimidine
 CC nucleotide (1). (1) is a catalytic DNA molecule, capable of hydrolytic
 CC cleavage of a nucleic acid phosphoester bond. The catalytic DNA molecule
 CC is used to cleave RNA of almost any sequence. This is useful for cleavage
 CC of single-stranded nucleic acid in the absence of a restriction
 CC endonuclease site at a specific position. This is useful in methods of
 CC e.g. cloning and genetic engineering. The catalytic core of the smallest
 CC enzyme is composed of only 12 residues, making this one of the smallest
 CC nucleic acid catalysts known. The catalytic core forms a compact hairpin
 CC structure displaying the 3 imidazole-containing residues. The enzyme can
 CC be made to cleave RNA of almost any sequence by simple alteration of the
 CC two substrate-recognition domains that surround the catalytic core. The
 CC enzyme operates with multiple turnover in the presence of micromolar
 CC concentrations of Zn²⁺, exhibiting saturation kinetics and a catalytic
 CC rate enhancement of approximately 1000000-fold compared to the
 CC uncatalysed reaction. The imidazole-containing DNA enzyme combines the
 CC substrate-recognition properties of nucleic acid enzymes and the chemical
 CC functionality of protein enzymes in a molecule that is small in size, yet
 CC versatile and catalytically efficient. The present sequence represents an
 CC oligonucleotide sequence which is used in the exemplification of the
 CC present invention
 XX
 XX Sequence 19 BP; 7 A; 2 C; 6 G; 4 T; 0 U; 0 Other;
 SQ
 Query Match 90.0%; Score 18; DB 3; Length 19;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 TCACCTATAGGAGAGATG 20
 DB 1 TCACCTATAGGAGAGATG 18
 RESULT 35
 AAF60641/c
 ID AAF60641 standard; DNA; 103 BP.
 XX
 AC AAF60641;
 XX
 DT 27-APR-2001 (first entry)
 XX
 DE Probe #13.
 XX
 KW T7 RNA polymerase promoter; probe; ss.
 XX
 OS Unidentified.
 XX
 XX WO200109377-A1.
 PN
 XX 08-FEB-2001.
 PD
 XX 31-JUL-2000; 2000WO-GB002962.
 PF
 XX 29-JUL-1999; 99GB-00017813.
 PR
 XX 17-AUG-1999; 99US-0149176P.
 XX
 XX (CYTO-) CYTOCELL LTD.
 XX
 XX Lloyd JS, Weston A, Cady DLN;
 XX
 XX WPI, 2001-182976/18.
 DR
 XX New complex formed by a hybridization reaction, useful for detecting a
 PT nucleic acid sequence of interest in sample, comprises the target nucleic

```

PT acid molecule and 2 or 3 probes.
PS Example 6; Page 27; 59pp; English.
CC The present invention relates to a complex formed by the hybridisation of
CC a target nucleic acid (e.g. a bacteriophage RNA polymerase promoter
CC sequence) and probe. The probes are useful in an assay for detecting the
CC presence of a nucleic acid sequence of interest, in a sample. The present
CC sequence is a probe which was used in the present invention
XX
XX Sequence 103 BP; 22 A; 26 C; 19 G; 36 T; 0 U; 0 Other;
SQ
Query Match 90.0%; Score 18; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGA 18
DB 54 ACTCACTATAGGAAGAGA 37
RESULT 36
AAP60585/c
ID AAP60585 standard; DNA; 109 BP.
AC AAP60585;
XX
XX 27-APR-2001 (first entry)
DE Probe #3.
XX RNA polymerase promoter; probe; disease marker detection; ss.
XX Unidentified.
XX WO200109376-A1.
XX 08-FEB-2001.
XX 31-JUL-2000; 2000WO-G8002946.
XX 29-JUL-1999; 99GB-00017816.
XX (CYTO-) CYTOCELL LTD.
XX Lloyd JS, Weston A, Cardy DLN, Marsh P;
XX WPI; 2001-182975/18.
XX Novel probe molecule useful for detecting nucleic acid sequence in a
XX sample, comprises single stranded nucleic acid sequence complementary to
XX target sequence, RNA polymerase promoter sequence, and blocking group.
XX Example 1; Page 14; 59pp; English.
XX The present invention relates to probes complementary to a single strand
XX of an RNA polymerase promoter sequence, and a blocking group adjacent or
XX substantially adjacent to the promoter sequence. The probes are useful
XX for detecting nucleic acid sequence of interest in a sample, for e.g. a
XX nucleic acid which is a marker of genetic or infectious disease. The
XX present sequence is one such probe which was used in the present
XX invention
XX Sequence 109 BP; 13 A; 37 C; 18 G; 41 T; 0 U; 0 Other;
XX
XX Query Match 90.0%; Score 18; DB 5; Length 109;
XX Best Local Similarity 100.0%; Pred. No. 35;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGA 18
DB 54 ACTCACTATAGGAAGAGA 37

```

```

RESULT 37
AAP60590/c
ID AAP60590 standard; DNA; 111 BP.
XX
XX AAP60590;
AC AAP60590;
XX
XX 27-APR-2001 (first entry)
DT Probe #5.
DB
XX RNA polymerase promoter; probe; disease marker detection; ss.
XX Unidentified.
XX WO200109376-A1.
XX 08-FEB-2001.
XX 31-JUL-2000; 2000WO-G8002946.
XX 29-JUL-1999; 99GB-00017816.
XX (CYTO-) CYTOCELL LTD.
XX Lloyd JS, Weston A, Cardy DLN, Marsh P;
XX WPI; 2001-182975/18.
XX Novel probe molecule useful for detecting nucleic acid sequence in a
XX sample, comprises single stranded nucleic acid sequence complementary to
XX target sequence, RNA polymerase promoter sequence, and blocking group.
XX Example 2; Page 17; 59pp; English.
XX The present invention relates to probes complementary to a single strand
XX of an RNA polymerase promoter sequence, and a blocking group adjacent or
XX substantially adjacent to the promoter sequence. The probes are useful
XX for detecting nucleic acid sequence of interest in a sample, for e.g. a
XX nucleic acid which is a marker of genetic or infectious disease. The
XX present sequence is one such probe which was used in the present
XX invention
XX Sequence 111 BP; 15 A; 36 C; 20 G; 40 T; 0 U; 0 Other;
XX
XX Query Match 90.0%; Score 18; DB 5; Length 111;
XX Best Local Similarity 100.0%; Pred. No. 35;
XX Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGA 18
DB 54 ACTCACTATAGGAAGAGA 37
RESULT 38
AAP60604/c
ID AAP60604 standard; DNA; 111 BP.
XX
XX AAP60604;
AC AAP60604;
XX
XX 27-APR-2001 (first entry)
DT Probe #15.
DB
XX RNA polymerase promoter; probe; disease marker detection; ss.
XX Unidentified.
XX WO200109376-A1.
XX 08-FEB-2001.
XX 31-JUL-2000; 2000WO-G8002946.
XX

```

```

XX 29-JUL-1999; 99GB-00017816.
XX (CYTO-) CYTOCELL LTD.
XX Lloyd JS, Weston A, Cardy DLN, Marsh P;
XX WPI; 2001-182975/18.
XX Novel probe molecule useful for detecting nucleic acid sequence in a
XX sample, comprises single stranded nucleic acid sequence complementary to
XX target sequence, RNA polymerase promoter sequence, and blocking group.
XX Example 6; Page 28; 59pp; English.
XX The present invention relates to probes complementary to a single strand
XX of an RNA polymerase promoter sequence, and a blocking group adjacent or
XX substantially adjacent to the promoter sequence. The probes are useful
XX for detecting nucleic acid sequence of interest in a sample, for e.g. a
XX nucleic acid which is a marker of genetic or infectious disease. The
XX present sequence is one such probe which was used in the present
XX invention.
XX Sequence 111 BP; 12 A; 29 C; 29 G; 41 T; 0 U; 0 Other;
Query Match 90.0%; Score 18; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGA 18
DB 54 ACTCACTATAGGAGAGA 37
RESULT 39
AAF60596/c
ID AAF60596 standard; DNA; 111 BP.
XX AAF60596;
XX 27-APR-2001 (first entry)
XX Probe #9.
XX RNA polymerase promoter; probe; disease marker detection; ss.
XX Unidentified.
XX WO200109376-A1.
XX 08-FEB-2001.
XX 31-JUL-2000; 2000WO-GB002946.
XX 29-JUL-1999; 99GB-00017816.
XX (CYTO-) CYTOCELL LTD.
XX Lloyd JS, Weston A, Cardy DLN, Marsh P;
XX WPI; 2001-182975/18.
XX Novel probe molecule useful for detecting nucleic acid sequence in a
XX sample, comprises single stranded nucleic acid sequence complementary to
XX target sequence, RNA polymerase promoter sequence, and blocking group.
XX Example 3; Page 21; 59pp; English.
XX The present invention relates to probes complementary to a single strand
XX of an RNA polymerase promoter sequence, and a blocking group adjacent or
XX substantially adjacent to the promoter sequence. The probes are useful
XX for detecting nucleic acid sequence of interest in a sample, for e.g. a
XX nucleic acid which is a marker of genetic or infectious disease. The
XX present sequence is one such probe which was used in the present
XX invention.
XX Sequence 111 BP; 12 A; 29 C; 29 G; 41 T; 0 U; 0 Other;
Query Match 90.0%; Score 18; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGA 18
DB 54 ACTCACTATAGGAGAGA 37
RESULT 39
AAF60596/c
ID AAF60596 standard; DNA; 111 BP.
XX AAF60596;
XX 27-APR-2001 (first entry)
XX Probe #9.
XX RNA polymerase promoter; probe; disease marker detection; ss.
XX Unidentified.
XX WO200109376-A1.
XX 08-FEB-2001.
XX 31-JUL-2000; 2000WO-GB002946.
XX 29-JUL-1999; 99GB-00017816.
XX (CYTO-) CYTOCELL LTD.
XX Lloyd JS, Weston A, Cardy DLN, Marsh P;
XX WPI; 2001-182975/18.
XX Novel probe molecule useful for detecting nucleic acid sequence in a
XX sample, comprises single stranded nucleic acid sequence complementary to
XX target sequence, RNA polymerase promoter sequence, and blocking group.
XX Example 3; Page 21; 59pp; English.
XX The present invention relates to probes complementary to a single strand
XX of an RNA polymerase promoter sequence, and a blocking group adjacent or
XX substantially adjacent to the promoter sequence. The probes are useful
XX for detecting nucleic acid sequence of interest in a sample, for e.g. a
XX nucleic acid which is a marker of genetic or infectious disease. The
XX present sequence is one such probe which was used in the present
XX invention.
XX Sequence 111 BP; 12 A; 29 C; 29 G; 41 T; 0 U; 0 Other;
Query Match 90.0%; Score 18; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGA 18
DB 54 ACTCACTATAGGAGAGA 37

```

```

CC Present sequence is one such probe which was used in the present
CC invention
XX Sequence 111 BP; 13 A; 35 C; 21 G; 42 T; 0 U; 0 Other;
Query Match 90.0%; Score 18; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGA 18
DB 54 ACTCACTATAGGAGAGA 37
RESULT 40
AAF60594/c
ID AAF60594 standard; DNA; 111 BP.
XX AAF60594;
XX 27-APR-2001 (first entry)
XX Probe #8.
XX RNA polymerase promoter; probe; disease marker detection; ss.
XX Unidentified.
XX WO200109376-A1.
XX 08-FEB-2001.
XX 31-JUL-2000; 2000WO-GB002946.
XX 29-JUL-1999; 99GB-00017816.
XX (CYTO-) CYTOCELL LTD.
XX Lloyd JS, Weston A, Cardy DLN, Marsh P;
XX WPI; 2001-182975/18.
XX Novel probe molecule useful for detecting nucleic acid sequence in a
XX sample, comprises single stranded nucleic acid sequence complementary to
XX target sequence, RNA polymerase promoter sequence, and blocking group.
XX Example 3; Page 20; 59pp; English.
XX The present invention relates to probes complementary to a single strand
XX of an RNA polymerase promoter sequence, and a blocking group adjacent or
XX substantially adjacent to the promoter sequence. The probes are useful
XX for detecting nucleic acid sequence of interest in a sample, for e.g. a
XX nucleic acid which is a marker of genetic or infectious disease. The
XX present sequence is one such probe which was used in the present
XX invention.
XX Sequence 111 BP; 13 A; 35 C; 21 G; 42 T; 0 U; 0 Other;
Query Match 90.0%; Score 18; DB 5; Length 111;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ACTCACTATAGGAGAGA 18
DB 54 ACTCACTATAGGAGAGA 37

```

Search completed: May 24, 2004, 11:14:45
Job time : 124.019 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2004, 10:22:21 ; Search time 1104.53 Seconds
(without alignments)
540.723 Million cell updates/sec

Title: US-10-144-679-2

Perfect score: 20

Sequence: 1 actactataggagagatg 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estm.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_est1.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estoa.*

17: em_gsa_hum.*

18: em_gsa_inv.*

19: em_gsa_pln.*

20: em_gsa_vrt.*

21: em_gsa_fun.*

22: em_gsa_cam.*

23: em_gsa_mus.*

24: em_gsa_pro.*

25: em_gsa_fod.*

26: em_gsa_phg.*

27: em_gsa_vrl.*

28: gb_gsa1.*

29: gb_gsa2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	19	95.0	632	14	CD767851
2	17.4	87.0	459	14	CD071925
3	17.4	87.0	556	28	AQ430253
4	17	85.0	496	9	AA567751

C	5	16.8	84.0	55	14	CB261880
	6	16.8	84.0	90	12	BM778744
	7	16.8	84.0	101	12	BG899265
C	8	16.8	84.0	146	10	BB837087
	9	16.8	84.0	247	10	AA447247
	10	16.8	84.0	252	10	BB837170
	11	16.8	84.0	262	10	AW185818
C	12	16.8	84.0	284	10	AW792821
	13	16.8	84.0	293	12	BL451235
C	14	16.8	84.0	321	10	BF359459
	15	16.8	84.0	335	14	CD598048
C	16	16.8	84.0	355	10	AW972800
	17	16.8	84.0	356	12	BG897743
	18	16.8	84.0	409	12	BG927617
C	19	16.8	84.0	416	10	AW971346
C	20	16.8	84.0	421	10	AW973262
C	21	16.8	84.0	432	10	AW969923
C	22	16.8	84.0	434	12	BG927056
C	23	16.8	84.0	438	10	AW979098
C	24	16.8	84.0	439	10	AW979113
C	25	16.8	84.0	441	10	AW971980
C	26	16.8	84.0	442	10	AW969893
C	27	16.8	84.0	446	10	AW972775
C	28	16.8	84.0	459	10	AW974684
C	29	16.8	84.0	463	10	AW969896
C	30	16.8	84.0	469	12	BG927057
C	31	16.8	84.0	473	10	AW970102
C	32	16.8	84.0	476	10	AW970875
C	33	16.8	84.0	478	10	AW976505
C	34	16.8	84.0	480	10	AW971417
C	35	16.8	84.0	487	10	AW969872
C	36	16.8	84.0	495	10	AW973174
C	37	16.8	84.0	496	10	AW969860
C	38	16.8	84.0	497	10	AW973757
C	39	16.8	84.0	497	10	AW975254
C	40	16.8	84.0	499	10	AW979054
C	41	16.8	84.0	500	12	BP190847
C	42	16.8	84.0	501	10	AW971097
C	43	16.8	84.0	507	10	AW973121
C	44	16.8	84.0	510	10	AW973393
C	45	16.8	84.0	510	10	AW979057

ALIGNMENTS

RESULT 1
CD767851

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CD767851 632 bp mRNA linear EST 02-JUL-2003
AGENCOURT 14714597 NICHOL RH Test2 Macaca mulatta cDNA clone
IMAGE:6973482 5', mRNA sequence.

CD767851

CD767851.1 GI:32426353

EST.

Macaca mulatta (rhesus monkey)

Macaca mulatta

Sukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopithecoidea; Macaca.

1 (bases 1 to 632)

NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP).

Unpublished (1997)

Tumor gene index

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: CGAPba-y@mail.nih.gov

Tissue Procurement: Dr. Tony M. Plant

cDNA Library Preparation: Invitrogen Corp

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA sequencing by: Agencourt Bioscience Corporation

Clone distribution: NC1-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLC9370 row: 0 column: 17
 High quality sequence stop: 468.
 Location/Qualifiers

FEATURES

Source

1..632
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /db_xref="taxon:9544"
 /clone="IMAGE-6973482"
 /tissue_type="Testis"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NICHD_RN_Test3"
 /note="Organ: Testis; Vector: pDNR-L19; Site 1: Sfil (ggccattggcc); Site 2: Sfil (ggccctcgcc); Site 3: Sfil (ggccctcgcc); 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGCGCGCATG-3' (30)BN-3' (where B = A, C, G, or T). Average insert size 1.67 kb (range 0.8-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 95.0%; Score 19; DB 14; Length 632;
 Best Local Similarity 100.0%; Pred. No. 95;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ACTCACTATAGGAGAGAT 19
 |||||
 15 ACTCACTATAGGAGAGAT 33

RESULT 2

CD071925/c
 LOCUS
 DEFINITION
 MA2-0036G-V373-G02-U.B MA2-0036 Schistosoma mansoni cDNA clone
 MA2-0036G-V373-G02.B, mRNA sequence.
 CD071925
 EST.
 CD071925.1 GI:34622962

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Schistosoma mansoni
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
 Strigoida; Schistosomatoidea; Schistosomatidae; Schistosoma.
 1 (bases 1 to 459)
 Verjovski-Almeida, S., DeMarco, R., Martins, E.A.L., Guimaraes, P.E.M., Ojopi, E.P.B., Paquola, A.C.M., Piazza, J.P., Nishiyama, M.Y. Jr., Kitajima, J.P., Adamson, R.E., Ashton, P.D., Bonaldo, M.P., Coulson, P.S., Dillon, G.P., Faras, L.P., Gregorio, S.P., Ho, P.L., Leite, R.A., Malaquias, L.C.C., Marques, R.C.P., Miyasato, P.A., Nascimento, A.L.T.O., Oliveira, F.P., Reis, B.M., Ribeiro, M.A., Sa, R.G., Stukart, G.C., Soares, M.B., Gargioni, C., Kavaroti, T., Rodrigues, V., Madeira, A.M.B.N., Wilson, R.A., Nench, C.F.M., Serubal, J.C., Leite, L.C.C. and Dias-Neto, E.
 Transcriptional analysis of the acclimated human parasite Schistosoma mansoni
 Nat. Genet. 35 (2), 148-157 (2003)

TITLE

JOURNAL

MEDLINE

COMMENT

Contact: Dr. Sergio Verjovski-Almeida
 Departamento de Bioquímica
 Instituto de Química - Universidade de São Paulo
 Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 São Paulo - SP, Brasil
 Tel: +55-11-3091-2173
 Fax: +55-11-3091-2186
 Email: verjowski@usp.br
 This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL <http://bioinfo.iq.usp.br/schisto/>
 Plate: MA2-0036G-V373 row: 2 column: G.
 Location/Qualifiers

FEATURES

Source

1..459
 /organism="Schistosoma mansoni"
 /mol_type="mRNA"
 /db_xref="taxon:6183"
 /clone="MA2-0036G-V373-G02.B"
 /sex="mixed pool"
 /dev_stages="adult"
 /lab_host="Mus musculus"
 /clone_lib="MA2-0036"
 /note="Vector: pGEM T-easy"

ORIGIN

Query Match 87.0%; Score 17.4; DB 14; Length 459;
 Best Local Similarity 94.7%; Pred. No. 5.3e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 2 CTCCTATAGGAGAGATG 20
 |||||
 198 CTCCTATAGGAGAGATG 180

RESULT 3

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Query Match 87.0%; Score 17.4; DB 28; Length 556;
 Best Local Similarity 94.7%; Pred. No. 5.7e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 CTCACTATAGGAGAGATG 20
 |||||
 DB 22 CTCACTATAGGAGAGATG 4

RESULT 4
 AA567751/c
 LOCUS
 DEFINITION HL01719 Sprime HL Drosophila melanogaster head Bluescript
 Drosophila melanogaster cDNA clone HL01719 5prime, mRNA sequences.
 AA567751
 VERSION
 AA567751.1 GI:2700081
 KEYWORDS
 BST.
 SOURCE
 Drosophila melanogaster (fruit fly)
 ORGANISM
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE
 1 (bases 1 to 496)
 AUTHORS
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Teang, G.,
 Lewis, S. and Rubin, G.M.
 TITLE
 BDGP/HMI Drosophila EST Project
 JOURNAL
 Unpublished (2001)
 COMMENT
 On Dec 18, 1997 this sequence version replaced gi:2340533.
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
 Plate: 17 row: B column: 7
 High quality sequence stop: 322.

FEATURES
 source
 1..496
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="BDGP_EST:BDcln022638"
 /db_xref="taxon:7227"
 /clone="HL01719"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="SOLR"
 /clone_lib="HL Drosophila melanogaster head Bluescript"
 /note="Organ: head-brain & sensory organ; Vector:
 Bluescript SK; Site 1: EcoRI; Site 2: XhoI; Constructed
 using Stratagene ZAP-cDNA Synthesis kit. Oligo dT-primed
 and directionally cloned at EcoRI and XhoI in Bluescript
 SK(+/-)."

ORIGIN
 Query Match 85.0%; Score 17; DB 9; Length 496;
 Best Local Similarity 100.0%; Pred. No. 8.6e+02;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 CACTATAGGAGAGATG 20
 |||||
 DB 61 CACTATAGGAGAGATG 45

RESULT 5
 CB261880/c
 LOCUS
 DEFINITION 83-B8880-008-014-P21-pB12 MP12-ADIS-008 Arabidopsis thaliana cDNA
 clone MP12p767E21140 5-PRIME, mRNA sequence.
 CB261880
 VERSION
 CB261880.1 GI:32886653
 KEYWORDS
 EST.
 SOURCE
 Arabidopsis thaliana (chale cress)

ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsais.
 REFERENCE
 1 (bases 1 to 55)
 AUTHORS
 Schmid, K.J., Soerensen, T.R., Stracke, R., Torjek, O., Altmann, T.,
 Mitchell-Olds, T. and Weishaar, B.
 TITLE
 Large-scale identification and analysis of genome-wide
 single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
 Genoms Res. 13 (6), 1250-1257 (2003)
 23883390
 MEDLINE
 PUBMED
 12799357
 COMMENT
 Contact: Weishaar B
 ADIS DNA core facility at MPIZ
 Max-Planck-Institute for Plant Breeding Research
 Carl-von-Linne Weg 10, 50829 Koeln, Germany
 Fax: 00492315062851
 Email: weishaar@mpiz-koeln.mpg.de
 Insert Length: 55 Std Error: 0.00
 Plate: 14 row: F column: 21
 Seg primer: PB12; GGTGGCGCGCTCTAG.
 Location/Qualifiers
 1..55
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="C24"
 /db_xref="GABI:601149"
 /db_xref="taxon:3702"
 /clone="MP12p767P21140"
 /tissue_type="seedling"
 /dev_stage="few days old seedlings"
 /lab_host="B. coli XLI-Blue MRP"
 /clone_lib="MP12-ADIS-008"
 /note="Vector: pBluescript SK (-); Site 1: EcoRI; Site 2:
 XhoI; cDNA library from Arabidopsis thaliana, accession
 C24; seedling; Lambda ZAPII phage library was made at the
 Max-Planck-Institute of Molecular Plant Physiology, Golem,
 Germany and mass-excised at the Max-Planck-Institute for
 Plant Breeding Research, Cologne, Germany; cloning sites
 EcoRI-XhoI; Note: Sequencing granted in the context of the
 GABI Arabidopsis Verbund I: Genetic Diversity.
 'Establishment of high-efficiency SNP-based mapping tools
 and development of methods for genome-wide mutation
 detection' PI: Bernd Weishaar Sequence submission managed
 by RZPD/GABI-Primary database: <http://gabi.rzpd.de> This
 clone is available from RZPD; contact RZPD (clone@rzpd.de)
 for further information."

ORIGIN
 Query Match 84.0%; Score 16.8; DB 14; Length 55;
 Best Local Similarity 90.0%; Pred. No. 4.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 34 ACTCACTATAGGAGAGATG 15

RESULT 6
 BM778744
 LOCUS
 DEFINITION 90 bp mRNA linear EST 04-MAR-2002
 fx22a04.y1 Zebrafish C32 14 splice embryo Danio rerio cDNA clone
 IMAGE:5620495 5', mRNA sequence.
 BM778744
 ACCESSION
 BM778744
 VERSION
 BM778744.1 GI:19108358
 KEYWORDS
 EST.
 SOURCE
 Danio rerio (zebrafish)
 ORGANISM
 Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE
 1 (bases 1 to 90)
 AUTHORS
 Clark, M., Johnson, S.L., Lebrach, H., Lee, R., Li, F., Marra, M.,

High quality sequence start: 9
High quality sequence stop: 146.

FEATURES

Location/Qualifiers
1..146
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="FNO089"
/note="Organ: prostate normal; Vector: puc18; Site 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORHSTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 146;
Best Local Similarity 90.0%; Pred. No. 6.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
|||||
DB 100 ACTCACTATAGGAGAGATG 81

RESULT 9

AM447247 247 bp mRNA linear EST 25-APR-2001
DEFINITION 88590 MARC 18OV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AM447247
VERSION AM447247.1 GI:6989034

KEYWORDS EST.
SOURCE Bos taurus (cow)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 247)
Smith, R.P.L., Grosse, W.M., Preking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.R., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
Chitko-McKown, C.G., Partee, G., Holt, I., Karamycheva, S., Liang, P.,
Quackenbush, J., and Keefe, J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

JOURNAL

21180013
MEDLINE
PUBMED 11282978

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smith@email.mars.uga.edu

Single pass sequencing. Bases called and trimmed with phred
v0.980504.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATACCAT

BACKWARD: GTTTCCTGATCGACG

Plate: 63 row: P column: 19

Seq primer: ATTAGGTGACATATAG.

Location/Qualifiers

1..247

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/ad_host="DH10B"

/clone_lib="MARC 18OV"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 247;
Best Local Similarity 90.0%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
|||||
DB 162 ACTCACTATAGGAGAGATG 181

RESULT 10

BH837170 252 bp mRNA linear EST 21-SEP-2000
DEFINITION RC2-FNO089-160600-014-h07 FNO089 Homo sapiens cDNA, mRNA sequence.
ACCESSION BH837170
VERSION BH837170.1 GI:10269548

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 252)

AUTHORS

Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hara, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL

20202663

MEDLINE

10737800

PUBMED

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil.

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-RC2-FNO089-160

600-014-h07&t3=2000-06-16&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 31

High quality sequence stop: 252.

FEATURES

Location/Qualifiers

1..252

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="FNO089"

/note="Organ: prostate normal; Vector: puc18; Site 1:

SmaI; Site 2: SmaI; A mini-library was made by cloning

products derived from ORHSTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the pUC 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 252;
Best Local Similarity 90.0%; Pred. No. 8.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
|||||

Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

Query Match 84.0%; Score 16.8; DB 14; Length 335;
 Best Local Similarity 90.0%; Pred. No. 9.2e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCCTATAGGAGAGATG 20
 DB 310 ACTAACCATAGGAGAGATG 329

RESULT 16
 AM972800/c 355 bp mRNA linear EST 01-JUN-2000
 LOCUS
 DEFINITION EST384895 MAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.
 ACCESSION AM972800
 VERSION AM972800.1 GI:8162646
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (Bases 1 to 355)
 AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharwad, S., Gaspard, R., Gay, C., Holt, I. S., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray

JOURNAL Unpublished (2000)
 COMMENT Contact: John Quackenbush

The Institute for Genomic Research
 7112 Medical Center Dr., Rockville, MD 20850, USA
 Tel.: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 310

Seq primer: Forward.

FEATURES
 source
 1..355
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGL"
 /note="Vector: pBluescriptSKm"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 355;
 Best Local Similarity 90.0%; Pred. No. 9.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCCTATAGGAGAGATG 20
 DB 337 ACTCCTATAGGAGAGATG 318

RESULT 17
 BG997743 356 bp mRNA linear EST 06-NOV-2001
 LOCUS
 DEFINITION HOA17-1-512 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA, mRNA sequence.

ACCESSION BG997743
 VERSION BG997743.1 GI:14307932
 KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (Bases 1 to 356)
 AUTHORS Kumar, S., Connor, J. R., Dodds, R. A., Halsey, W., Van Horn, M., Mao, J., Sathya, G., Mui, P., Agarwal, P., Badger, A. M., Lee, J. C., Gowen, M. and Lark, M. W.

TITLE Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries

JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 PUBMED 11597177
 COMMENT Contact: Sanjay Kumar
 UW2109

GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay.kumar-logsk.com
 Seq primer: 17.

FEATURES
 source
 Location/Qualifiers
 1..356
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="cartilage"
 /lab_host="E.coli DH10 B"
 /clone_lib="HOA (Human Osteoarthritic Cartilage)"
 /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI; Directional"

ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 356;
 Best Local Similarity 90.0%; Pred. No. 9.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCCTATAGGAGAGATG 20
 DB 9 ACTCCTATAGGAGAGATG 28

RESULT 18
 BG927617 409 bp mRNA linear EST 06-NOV-2001
 LOCUS

DEFINITION HNC41-1-G12.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA sequence.

ACCESSION BG927617
 VERSION BG927617.1 GI:14322140
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (Bases 1 to 409)
 Kumar, S., Connor, J. R., Dodds, R. A., Halsey, W., Van Horn, M., Mao, J., Sathya, G., Mui, P., Agarwal, P., Badger, A. M., Lee, J. C., Gowen, M. and Lark, M. W.

TITLE Identification and initial characterization of 5000 expressed sequenced tags (ESTs) each from adult human normal and osteoarthritic cartilage cDNA libraries

JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 PUBMED 11597177

COMMENT Contact: Sanjay Kumar

GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay.kumar-logsk.com

Seq primer: 17.

FEATURES
 source
 Location/Qualifiers
 1..409
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="cartilage"
 /lab_host="E.coli DH10 B"
 /clone_lib="HNC (Human Normal Cartilage)"
 /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI; Directional"

ORIGIN

Query Match 84.0%; Score 16.8; DB 12; Length 409;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 16 ACTCACTATAGGAGAGATG 35

RESULT 19
 AW971346/c
 LOCUS AW971346 416 bp mRNA linear EST 01-JUN-2000
 DEFINITION EST383435 MAGS resequences, MAGL Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW971346
 VERSION AW971346.1 GI:8161191

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 416)
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
 Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
 Quackenbush, J.

REFERENCE
 AUTHORS Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 288

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..416
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGS resequences, MAGL"
 /note="Vector: pBluescriptSKm"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 416;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 388 ACTCACTATAGGAGAGATG 369

RESULT 20
 AW973262/c
 LOCUS AW973262 421 bp mRNA linear EST 01-JUN-2000
 DEFINITION EST385360 MAGS resequences, MAGM Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW973262
 VERSION AW973262.1 GI:8163120

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 421)
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
 Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
 Quackenbush, J.

REFERENCE
 AUTHORS Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 Contact: John Quackenbush

TITLE
 JOURNAL
 COMMENT

The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 316

Seq primer: Forward.
 Location/Qualifiers
 1..421
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGS resequences, MAGM"
 /note="Vector: pBluescriptSKm"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 421;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 381 ACTCACTATAGGAGAGATG 362

RESULT 21
 AW969923/c
 LOCUS AW969923 432 bp mRNA linear EST 01-JUN-2000
 DEFINITION EST382001 MAGS resequences, MAGK Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW969923
 VERSION AW969923.1 GI:8159767

KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 432)
 Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
 Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
 Quackenbush, J.

REFERENCE
 AUTHORS Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Unpublished (2000)
 Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 271

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..432
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGS resequences, MAGK"
 /note="Vector: pBluescriptSKm"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 432;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 368 ACTCACTATAGGAGAGATG 349

RESULT 22
 BG927056
 LOCUS BG927056 434 bp mRNA linear EST 06-NOV-2001
 DEFINITION HNC8-1-G10.R FMC (Human Normal Cartilage) Homo sapiens cDNA, mRNA

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..432
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGS resequences, MAGK"
 /note="Vector: pBluescriptSKm"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 432;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 368 ACTCACTATAGGAGAGATG 349

RESULT 22
 BG927056
 LOCUS BG927056 434 bp mRNA linear EST 06-NOV-2001
 DEFINITION HNC8-1-G10.R FMC (Human Normal Cartilage) Homo sapiens cDNA, mRNA

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..432
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGS resequences, MAGK"
 /note="Vector: pBluescriptSKm"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 432;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 368 ACTCACTATAGGAGAGATG 349

RESULT 22
 BG927056
 LOCUS BG927056 434 bp mRNA linear EST 06-NOV-2001
 DEFINITION HNC8-1-G10.R FMC (Human Normal Cartilage) Homo sapiens cDNA, mRNA

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 Location/Qualifiers
 1..432
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGS resequences, MAGK"
 /note="Vector: pBluescriptSKm"

ORIGIN
 Query Match 84.0%; Score 16.8; DB 10; Length 432;
 Best Local Similarity 90.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
 |||||
 DB 368 ACTCACTATAGGAGAGATG 349

RESULT 22
 BG927056
 LOCUS BG927056 434 bp mRNA linear EST 06-NOV-2001
 DEFINITION HNC8-1-G10.R FMC (Human Normal Cartilage) Homo sapiens cDNA, mRNA

TITLE
 JOURNAL
 COMMENT

```

sequence.
ACCESSION      BQ927056
VERSION        BQ927056.1  GI:14321579
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
REFERENCE      1 (bases 1 to 434)
AUTHORS       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               Kumar, S., Connor, J.R., Dodds, R.A., Halsey, M., Van Horn, M., Mao, J.,
               Sathya, G., Muir, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M. and
               Lark, M.W.
TITLE         Identification and initial characterization of 5000 expressed
               sequenced tags (ESTs) each from adult human normal and
               osteoarthritic cartilage cDNA libraries
JOURNAL        Osteoarthr. Cartil. 9 (7), 641-653 (2001)
MEDLINE        21482651
PubMed        11597177
COMMENT       Contact: Sanjay Kumar
               UN2109
               GlaxoSmithKline
               709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
               Tel: 610-270-7245
               Fax: 610-270-5598
               Email: sanjay.kumar-legsk.com
               Seq primer: 17.
FEATURES       Location/Qualifiers
               1..434
               /organism="Homo sapiens"
               /mol_type="cDNA"
               /db_xref="taxon:9606"
               /tissue_type="cartilage"
               /lab_host="H. coli DH10 B"
               /clone_lib="HNC (Human Normal Cartilage)"
               /note="Vector: pSPORT 1; Site 1: SalI; Site 2: NotI;
               Directional"
ORIGIN
Query Match      84.0%; Score 16.8; DB 12; Length 434;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY      1 ACTCACTATAGGAGAGATG 20
DB      13 ACTCACTATAGGAGAGCTG 32
RESULT 23
AW979098/c
LOCUS           AW979098      438 bp      mRNA      linear      EST 02-JUN-2000
DEFINITION      EST391208 MAGE resequences, MAGP Homo sapiens cDNA, mRNA sequence.
ACCESSION       AW979098
VERSION         AW979098.1  GI:8170383
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
REFERENCE       1 (bases 1 to 438)
AUTHORS         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
               Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C.,
               Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
               Quackenbush, J.
TITLE           Assessment of gene expression patterns in a model of colon tumor
               metastasis using a 19,200 element cDNA microarray
JOURNAL         Unpublished (2000)
COMMENT         Contact: John Quackenbush
               The Institute for Genomic Research
               9712 Medical Center Dr., Rockville, MD 20850, USA
               Tel: 301 838 3528
               Fax: 301 838 0208
               Email: johnq@tigr.org
               Plate: 402
               Seq primer: Forward.
FEATURES       Location/Qualifiers
               1..438
               /organism="Homo sapiens"
               /mol_type="cDNA"
               /db_xref="taxon:9606"
               /clone_lib="MAGE resequences, MAGP"
               /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 439;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY      1 ACTCACTATAGGAGAGATG 20
DB      399 ACTCACTATAGGAGAGCTG 380
RESULT 25
AW971980/c
LOCUS           AW971980      441 bp      mRNA      linear      EST 01-JUN-2000
DEFINITION      EST384189 MAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION       AW971980
VERSION         AW971980.1  GI:8161946
KEYWORDS        EST.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
REFERENCE       1 (bases 1 to 441)
AUTHORS         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
AUTHORS      Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
              Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
              Quackenbush,J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 300
              Seq primer: Forward.
              Location/Qualifiers
FEATURES     source
              1..441
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="MAGE resequencences, MAGL"
              /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 441;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 ACTCACTATAGGAAAGGATG 20
Db      429 ACTCACTATAGGAAAGGCTG 410

RESULT 26
AW969893/c
LOCUS      AW969893
DEFINITION EST1381971 MAGE resequencences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW969893
VERSION     AW969893.1 GI:8159737
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 442)
AUTHORS      Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
              Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
              Quackenbush,J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 270
              Seq primer: Forward.
              Location/Qualifiers
FEATURES     source
              1..442
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="MAGE resequencences, MAGK"
              /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 442;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 ACTCACTATAGGAAAGGATG 20
Db      429 ACTCACTATAGGAAAGGCTG 410

REFERENCE
AUTHORS      Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
              Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
              Quackenbush,J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 270
              Seq primer: Forward.
              Location/Qualifiers
FEATURES     source
              1..442
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="MAGE resequencences, MAGK"
              /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 442;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 ACTCACTATAGGAAAGGATG 20
Db      429 ACTCACTATAGGAAAGGCTG 410

```

```

Db      428 ACTCACTATAGGAAAGGTTG 409

RESULT 27
AW972775/c
LOCUS      AW972775
DEFINITION EST1384869 MAGE resequencences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW972775
VERSION     AW972775.1 GI:8162621
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 446)
AUTHORS      Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
              Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
              Quackenbush,J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 310
              Seq primer: Forward.
              Location/Qualifiers
FEATURES     source
              1..446
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone_lib="MAGE resequencences, MAGL"
              /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 446;
Best Local Similarity 90.0%; Pred. No. 1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      1 ACTCACTATAGGAAAGGATG 20
Db      417 ACTCACTATAGGAAAGGCTG 398

RESULT 28
AW974684/c
LOCUS      AW974684
DEFINITION EST1385773 MAGE resequencences, MAGM Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW974684
VERSION     AW974684.1 GI:8165871
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 459)
AUTHORS      Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
              Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
              Quackenbush,J.
TITLE        Assessment of gene expression patterns in a model of colon tumor
              metastasis using a 19,200 element cDNA microarray
JOURNAL      Unpublished (2000)
COMMENT      Contact: John Quackenbush
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 3528
              Fax: 301 838 0208
              Email: johnq@tigr.org
              Plate: 337

```

```

FEATURES
  source
    Seq primer: Forward.
    Location/Qualifiers
      1..459
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone_lib="MAGE resequences, MAGE"
        /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 459;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGCTG 20
    |||||
DB 442 ACTCACTATAGGAAGAGCTG 423
    |||||
RESULT 29
LOCUS AW969896/c 463 bp mRNA linear EST 01-JUN-2000
DEFINITION EST181974 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW969896
VERSION AW969896.1 GI:8159740
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 463)
  Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J.,
  Sathe, G., Mul, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M., and
  Izar, M.W.
  Identification and initial characterization of 5000 expressed
  sequenced tags (ESTs) each from adult human normal and
  osteoarthritic cartilage cDNA libraries
  Osteoarthr. Cartil. 9 (7), 641-653 (2001)
  21482651
  11597177
  Contact: Sanjay Kumar
  UW2109
  GlaxoSmithKline
  709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
  Tel: 610-270-7245
  Fax: 610-270-5538
  Email: sanjay_kumar-legsk.com
  Seq primer: F7.
  Location/Qualifiers
    1..469
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /tissue_type="cartilage"
      /lab_hosts="E.coli DH10 B"
      /clone_lib="HNC (Human Normal Cartilage)"
      /notes="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
      Directional"
ORIGIN
Query Match      84.0%; Score 16.8; DB 12; Length 469;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGCTG 20
    |||||
DB 14 ACTCACTATAGGAAGAGCTG 33
    |||||
RESULT 31
LOCUS AW970102/c 473 bp mRNA linear EST 01-JUN-2000
DEFINITION EST382183 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW970102
VERSION AW970102.1 GI:8159947
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 473)
  Hegde, P., Qi, R., Abernathy, K., Charap, S., Caspard, R., Gay, C.,
  Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
  Quackenbush, J.
  Assessment of gene expression patterns in a model of colon tumor
  metastasis using a 19,200 element cDNA microarray
  Unpublished (2000)
  Contact: John Quackenbush
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 3528
  Fax: 301 838 0208
  Email: jchaq@igr.org
  Plate: 270
  Seq primer: Forward.
  Location/Qualifiers
    1..463
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone_lib="MAGE resequences, MAGE"
      /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 463;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGCTG 20
    |||||
DB 448 ACTCACTATAGGAAGAGCTG 429
    |||||
RESULT 30
LOCUS BG927057 469 bp mRNA linear EST 06-NOV-2001
DEFINITION NC88-I-G12.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
sequence.
ACCESSION BG927057
VERSION BG927057.1 GI:14321580
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
  AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 469)
  Kumar, S., Connor, J.R., Dodds, R.A., Halsey, W., Van Horn, M., Mao, J.,
  Sathe, G., Mul, P., Agarwal, P., Badger, A.M., Lee, J.C., Gowen, M., and
  Izar, M.W.
  Identification and initial characterization of 5000 expressed
  sequenced tags (ESTs) each from adult human normal and
  osteoarthritic cartilage cDNA libraries
  Osteoarthr. Cartil. 9 (7), 641-653 (2001)
  21482651
  11597177
  Contact: Sanjay Kumar
  UW2109
  GlaxoSmithKline
  709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
  Tel: 610-270-7245
  Fax: 610-270-5538
  Email: sanjay_kumar-legsk.com
  Seq primer: F7.
  Location/Qualifiers
    1..469
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /tissue_type="cartilage"
      /lab_hosts="E.coli DH10 B"
      /clone_lib="HNC (Human Normal Cartilage)"
      /notes="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
      Directional"
ORIGIN
Query Match      84.0%; Score 16.8; DB 12; Length 469;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ACTCACTATAGGAAGAGCTG 20
    |||||
DB 14 ACTCACTATAGGAAGAGCTG 33
    |||||
RESULT 31
LOCUS AW970102/c 473 bp mRNA linear EST 01-JUN-2000
DEFINITION EST382183 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.
ACCESSION AW970102
VERSION AW970102.1 GI:8159947
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
  AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 473)
  Hegde, P., Qi, R., Abernathy, K., Charap, S., Caspard, R., Gay, C.,
  Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
  Quackenbush, J.
  Assessment of gene expression patterns in a model of colon tumor
  metastasis using a 19,200 element cDNA microarray
  Unpublished (2000)
  Contact: John Quackenbush
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 3528
  Fax: 301 838 0208
  Email: jchaq@igr.org
  Plate: 273
  Seq primer: Forward.
  Location/Qualifiers
    1..473
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone_lib="MAGE resequences, MAGE"

```



```

ORIGIN
/note="Vector: pBluescriptSKM"

Query Match      84.0%; Score 16.8; DB 10; Length 473;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
Db 444 ACTCACTATAGGAGAGCTG 425

RESULT 32
AW970875/C
LOCUS      476 bp mRNA linear EST 01-JUN-2000
DEFINITION EST382958 MAGS resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW970875
VERSION     AW970875.1 GI:8160720
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 476)
AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
           Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
           Quackenbush,J.
TITLE     Assessment of gene expression patterns in a model of colon tumor
           metastasis using a 19,200 element cDNA microarray
JOURNAL
COMMENT   Contact: John Quackenbush
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 3528
           Fax: 301 838 0208
           Email: johnq@tigr.org
           Plate: 289
           Seq primer: Forward.
           Location/Qualifiers
               1..476
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone_lib="MAGS resequences, MAGN"
               /note="Vector: pBluescriptSKM"

ORIGIN
/note="Vector: pBluescriptSKM"

Query Match      84.0%; Score 16.8; DB 10; Length 478;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
Db 462 ACTCACTATAGGAGAGCTG 443

RESULT 33
AW976505/C
LOCUS      478 bp mRNA linear EST 02-JUN-2000
DEFINITION EST388614 MAGS resequences, MAGN Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW976505
VERSION     AW976505.1 GI:8167734
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 478)
AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
           Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
           Quackenbush,J.
TITLE     Assessment of gene expression patterns in a model of colon tumor
           metastasis using a 19,200 element cDNA microarray
JOURNAL
COMMENT   Contact: John Quackenbush
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 3528
           Fax: 301 838 0208
           Email: johnq@tigr.org
           Plate: 289
           Seq primer: Forward.
           Location/Qualifiers
               1..480
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone_lib="MAGS resequences, MAGL"
               /note="Vector: pBluescriptSKM"

ORIGIN
/note="Vector: pBluescriptSKM"

Query Match      84.0%; Score 16.8; DB 10; Length 480;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
Db 427 ACTCACTATAGGAGAGCTG 408

RESULT 35
AW969872/C

```

```

JOURNAL
COMMENT   Unpublished (2000)
           Contact: John Quackenbush
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 3528
           Fax: 301 838 0208
           Email: johnq@tigr.org
           Plate: 362
           Seq primer: Forward.
           Location/Qualifiers
               1..478
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone_lib="MAGS resequences, MAGN"
               /note="Vector: pBluescriptSKM"

ORIGIN
/note="Vector: pBluescriptSKM"

Query Match      84.0%; Score 16.8; DB 10; Length 478;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
Db 461 ACTCACTATAGGAGAGCTG 442

RESULT 34
AW971417/C
LOCUS      480 bp mRNA linear EST 01-JUN-2000
DEFINITION EST383506 MAGS resequences, MAGL Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW971417
VERSION     AW971417.1 GI:8161262
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 480)
AUTHORS   Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
           Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
           Quackenbush,J.
TITLE     Assessment of gene expression patterns in a model of colon tumor
           metastasis using a 19,200 element cDNA microarray
JOURNAL
COMMENT   Contact: John Quackenbush
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 3528
           Fax: 301 838 0208
           Email: johnq@tigr.org
           Plate: 289
           Seq primer: Forward.
           Location/Qualifiers
               1..480
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone_lib="MAGS resequences, MAGL"
               /note="Vector: pBluescriptSKM"

ORIGIN
/note="Vector: pBluescriptSKM"

Query Match      84.0%; Score 16.8; DB 10; Length 480;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAGAGATG 20
    |||||
Db 427 ACTCACTATAGGAGAGCTG 408

RESULT 35
AW969872/C

```



```

LOCUS      AW969872      487 bp      mRNA      linear      EST 01-JUN-2000
DEFINITION EST381950 MAGK resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW969872
VERSION     AW969872.1 GI:8159716
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 487)
AUTHORS     Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
            Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
            Quackenbush, J.
TITLE       Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
JOURNAL     Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 270
            Seq primer: Forward.
FEATURES    source
            Location/Qualifiers
            1..487
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_lib="MAGK resequences, MAGK"
            /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 487;
Best Local Similarity 90.0%; Pred. NO. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACTCACTATAGGAGAGATG 20
        |||
        467 ACTCACTATAGGAGAGCTG 448

RESULT 36
AW973174/c
LOCUS      AW973174      495 bp      mRNA      linear      EST 01-JUN-2000
DEFINITION EST385272 MAGK resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW973174
VERSION     AW973174.1 GI:8163020
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 495)
AUTHORS     Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
            Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
            Quackenbush, J.
TITLE       Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
JOURNAL     Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 315
            Seq primer: Forward.
FEATURES    source
            Location/Qualifiers
            1..495
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_lib="MAGK resequences, MAGK"
            /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 496;
Best Local Similarity 90.0%; Pred. NO. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACTCACTATAGGAGAGATG 20
        |||
        495 ACTCACTATAGGAGAGTTG 476

RESULT 37
AW969860/c
LOCUS      AW969860      496 bp      mRNA      linear      EST 01-JUN-2000
DEFINITION EST381938 MAGK resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW969860
VERSION     AW969860.1 GI:8159704
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 496)
AUTHORS     Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
            Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
            Quackenbush, J.
TITLE       Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
JOURNAL     Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 270
            Seq primer: Forward.
FEATURES    source
            Location/Qualifiers
            1..496
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_lib="MAGK resequences, MAGK"
            /note="Vector: pBluescriptSKm"
ORIGIN
Query Match      84.0%; Score 16.8; DB 10; Length 496;
Best Local Similarity 90.0%; Pred. NO. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  ACTCACTATAGGAGAGATG 20
        |||
        495 ACTCACTATAGGAGAGTTG 476

RESULT 38
AW973757/c
LOCUS      AW973757      497 bp      mRNA      linear      EST 02-JUN-2000
DEFINITION EST385858 MAGK resequences, MAGK Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW973757
VERSION     AW973757.1 GI:8164942
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 497)
AUTHORS     Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,
            Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
            Quackenbush, J.
TITLE       Assessment of gene expression patterns in a model of colon tumor
            metastasis using a 19,200 element cDNA microarray
JOURNAL     Unpublished (2000)
COMMENT     Contact: John Quackenbush
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 3528
            Fax: 301 838 0208
            Email: johnq@tigr.org
            Plate: 315
            Seq primer: Forward.
FEATURES    source
            Location/Qualifiers
            1..497
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_lib="MAGK resequences, MAGK"
            /note="Vector: pBluescriptSKm"
ORIGIN

```

metastasis using a 19,200 element cDNA microarray

JOURNAL COMMENT

Unpublished (2000)
Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnqu@tigr.org
Plate: 324

Seq primer: Forward.

FEATURES source

Location/Qualifiers
1..497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="MAGE resequencences, MAGM"
/notes="Vector: pBluescriptSKm"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 497;
Best Local Similarity 90.0%; Pred. No. 1.1e+03;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGTG 20

DB 483 ACTCACTATAGGAAGAGTG 464

RESULT 39

AW975254/c

LOCUS

EST387362 MAGE resequencences, MAGM Homo sapiens cDNA, mRNA sequence.

DEFINITION

AW975254

ACCESSION

AW975254.1

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,

Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeaman, T. J. and

Quackenbush, J.

Assessment of gene expression patterns in a model of colon tumor

metastasis using a 19,200 element cDNA microarray

Unpublished (2000)

Contact: John Quackenbush

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 3528

Fax: 301 838 0208

Email: johnqu@tigr.org

Plate: 346

Seq primer: Forward.

Location/Qualifiers

1..497

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="MAGE resequencences, MAGM"

/notes="Vector: pBluescriptSKm"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 497;

Best Local Similarity 90.0%; Pred. No. 1.1e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGTG 20

DB 480 ACTCACTATAGGAAGAGTG 461

RESULT 40

AW979054/c

LOCUS

DEFINITION

AW979054

ACCESSION

AW979054.1

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,

Holt, I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeaman, T. J. and

Quackenbush, J.

Assessment of gene expression patterns in a model of colon tumor

metastasis using a 19,200 element cDNA microarray

Unpublished (2000)

Contact: John Quackenbush

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 3528

Fax: 301 838 0208

Email: johnqu@tigr.org

Plate: 402

Seq primer: Forward.

Location/Qualifiers

1..499

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone_lib="MAGE resequencences, MAGM"

/notes="Vector: pBluescriptSKm"

ORIGIN

Query Match 84.0%; Score 16.8; DB 10; Length 499;

Best Local Similarity 90.0%; Pred. No. 1.1e+03;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTCACTATAGGAAGAGTG 20

DB 457 ACTCACTATAGGAAGAGTG 438

Search completed: May 24, 2004, 12:29:52

Job time : 1107.53 secs